

# THÈSE

présentée

A LA FACULTÉ DES SCIENCES  
DE L'UNIVERSITÉ DE PARIS

pour obtenir

LE TITRE DE DOCTEUR DE 3ème CYCLE

Spécialité : STATISTIQUES MATHÉMATIQUES

par

Régine RÉNÉ-CHAUME

*Attaché de Recherches de l'O. R. S. T. O. M.*

---

ETUDE DE LA STRUCTURE DE LA VARIANCE GÉNÉTIQUE  
D'UNE POPULATION PANMICTIQUE À L'ÉQUILIBRE  
CHEZ UNE ESPÈCE AUTOTÉTAPLOÏDE

---

Soutenue le

devant la Commission d'Examen

Jury { MM. D. DUGUÉ      Président  
          Y. DEMARLY } Examineurs  
          G. PRÉVOST }

O. R. S. T. O. M.  
PARIS  
1968

T H E S E

présentée à

LA FACULTE DES SCIENCES DE PARIS

pour l'obtention

du DOCTORAT 3<sup>e</sup> CYCLE

Spécialité : STATISTIQUES

Mention : STATISTIQUES MATHÉMATIQUES

par Régine RÈNÉ-CHAUME

Sujet de la Thèse : Etude de la structure de la variance génétique d'une  
population panmictique à l'équilibre chez une espèce  
autotétraploïde.

Soutenue le

devant la Commission

composée de : Président D.DUGUÉ

Examineurs Y.DEMARLY

G.PREVOST

Invité A.GALLAIS

AVANT-PROPOS

Avant de présenter les résultats de ce travail, je voudrais exprimer ici ma profonde et respectueuse gratitude à toutes les personnes qui m'ont donné la possibilité de mener à bien cette étude. Mes sentiments très reconnaissants vont particulièrement à celles qui ont bien voulu accepter de constituer mon Jury.

Je remercie vivement Monsieur le Professeur DUGUE, Directeur de l'Institut de Statistique de l'Université de Paris, qui par son enseignement m'a initiée aux Statistiques Mathématiques et qui a accepté la Présidence du Jury.

Ma profonde reconnaissance va à Monsieur DEMARLY, Directeur de Recherches à l'Institut National de la Recherche Agronomique, auquel je dois le choix d'un sujet qui m'a fait découvrir et aimer la génétique quantitative. Que Monsieur le Directeur Général de l'Institut National de la Recherche Agronomique trouve ici l'expression de ma respectueuse gratitude pour m'avoir donné accès aux laboratoires de la Station d'Amélioration des Plantes Fourragères de Lusignan. Monsieur DEMARLY m'y a accueillie avec beaucoup de bienveillance, je le remercie d'avoir tout fait pour me faciliter la tâche.

J'y ai trouvé également Monsieur GALLAIS, Chargé de Recherches, qui par ses nombreux conseils et par son aide permanente lors de mes séjours, m'a aimablement guidée dans la réalisation de ce travail, qu'il trouve ici l'expression de toute ma gratitude.

Monsieur GUY, Chargé de Recherches, et le personnel du laboratoire de la luzerne m'ont apporté une collaboration technique dans la création du matériel expérimental et la mise en place des essais je les en remercie bien vivement.

### III

Mes remerciements reviennent également à Monsieur DEJARDIN, Maître de Recherches à l'Office de la Recherche Scientifique et Technique Outre-Mer, qui m'a prodigué tout au long de ce travail de nombreux conseils et m'a guidée dans toutes les interprétations statistiques.

Je remercie Monsieur le Professeur PREVOST de m'avoir reçue avec beaucoup de bienveillance et de s'être intéressé aux résultats de mon étude.

Je prie Monsieur le Professeur CAMUS, Directeur Général de l'Office de la Recherche Scientifique et Technique Outre-Mer de trouver ici l'expression de ma profonde reconnaissance pour m'avoir autorisée à entreprendre ce travail.

Mes remerciements vont enfin au Service de Documentation de l'O.R.S.T.O.M. pour la mise en page de ce mémoire et à Madame ANDRE s'est chargée de la dactylographie.

P L A NINTRODUCTION1ère partie : Théorie de l'analyse diallèle chez les tétraploïdes

## I - Définitions des croisements diallèles

## II - Analyse statistique

## A - Analyse de la variance "entre hybrides"

## 1°) Généralités

## 2°) Modèle mathématique

## B - Etude des aptitudes à la combinaison

## 1°) Généralités

## 2°) Modèle I

## 3°) Modèle II

## C - Décomposition des effets réciproques

## D - Analyse de la variance "entre et intra parcelles"

## E - Analyse des autofécondations

## III - Analyse génétique

## A - Hypothèses

B - Calcul des espérances mathématiques des carrés moyens en  
fonction du modèle génétique

## C - Définitions des variances d'additivité, de dominance, d'épistasie

## 1°) Cas d'un locus

## 2°) Cas de n loci

D - Décomposition des covariances entre individus apparentés en  
fonction des composants de la variance génétique

## 1°) Cas d'un locus

2°) Cas de deux loci

E - Calcul des coefficients ( $\phi$ ,  $\psi$ ,  $\phi'$ ,  $\psi'$ ) de parenté

F - Etude d'une population tétraploïde après une génération  
d'autofécondations

1°) Moyenne de la population

2°) Variance de la population

3°) Covariance entre individus apparentés

2ème partie : Application à l'étude d'une population allogame autotétraploïde :

la luzerne du Puits

I - Principes

1°) Etude des aptitudes à la combinaison

2°) Etude de la variabilité des effets géniques

II - Matériel et méthodes

1°) Choix de la population

2°) Optention du matériel

3°) Plan d'expérience

4°) Caractères étudiés

III - Analyse des résultats

1°) Etude des hybrides

2°) Etude des clones

3°) Etude des autofécondations

Conclusion

Bibliographie

Annexes

### Introduction

En amélioration des plantes du point de vue de la sélection, l'intérêt réside dans l'étude des caractères quantitatifs au sein des populations pour apporter des informations sur le mode moyen d'action des gènes et leur variabilité. Le phénotype d'un individu est le résultat des effets héréditaires modifiés par le milieu. En restreignant cette définition au cas où les phénomènes sont additifs on peut écrire que la valeur phénotypique d'un caractère (P) est la résultante d'une valeur génotypique (G) et des effets d'environnement ;  
d'où l'expression :

$$P = G + E$$

En faisant l'hypothèse d'indépendance entre le génotype et l'environnement les variances  $\sigma_G^2$  (génétique) et  $\sigma_E^2$  (environnement) s'additionnent :  
la valeur génotypique dépend des effets des gènes qui se manifestent :

A) directement : effet moyen d'additivité (A)

B) indirectement : a) effet moyen d'interaction entre allèles à un même locus ou effet de dominance (D)

b) effet moyen d'interaction entre allèles à des loci différents ou effet d'épistasie (I)

d'où en faisant l'hypothèse de l'additivité des effets l'expression de la valeur génotypique :

$$G = A + D + I$$

La connaissance de l'importance relative des variances d'additivité, dominance, épistasie servira de base à la construction d'un schéma de sélection efficace, pour un caractère quantitatif donné, chez les plantes allogames autotétraploïdes, comme elle l'est déjà chez les plantes allogames diploïdes.



Dans la sélection pour la vigueur il s'agit d'utiliser au maximum les phénomènes d'hétérosis. Cet aspect est d'autant plus important chez les plantes autotétraploïdes qu'il existe différents niveaux d'hétérozygotie possibles à un même locus, dans l'hypothèse de multiallélisme. Ainsi sans attacher d'importance à la nature des allèles présents cinq états peuvent être rencontrés à un locus : monogénique, digénique (simplex, duplex) trigénique et tétragénique.

Mais compte tenu des modalités de réalisation de la valeur génotypique, l'état d'un locus ne peut pas à priori être considéré comme déterminant. La valeur d'un génotype dépend de son niveau d'hétérozygotie, mais aussi des allèles présents, de leur action moyenne directe (effet d'additivité) et de leurs interactions avec le reste du génome (effets de dominance, épistasie).

Le progrès de la sélection sera donc lié à l'efficacité du choix des allèles les plus efficaces dans le génome. Ce progrès dépendra donc de l'importance et de la variabilité des effets géniques à un locus et des interactions de ces effets entre loci. Plus un effet favorable en moyenne est variable, plus le progrès attendu sera élevé.

Il importe donc d'apprécier la variabilité des effets géniques avec un modèle approprié. Or chez les espèces reconnues autopolyploïdes, les schémas d'analyse de la variance génétique des diploïdes ont été appliqués tels quels sans se soucier des particularités dues à l'état polyploïde.

Chez les espèces diploïdes différents modèles d'analyse de la variance génétique d'une population à l'équilibre ont été proposés :

Expériences de types I et II de Comstock et Robinson (1952)

Analyse diallèle développée par Hayman (1954 a et b) pour les populations autogames, bialléliques, généralisée par Kempthorne et Griffing (1956 a et b).

Dans l'expérience de type I de Comstock et Robinson,  $m$  individus mâles pris au hasard sont croisés chacun avec un échantillon au hasard de  $F$  individus femelles ; chaque descendance est testée dans  $r$  répétitions. Ce schéma est surtout valable pour l'étude des populations animales. Il peut être modifié (expérience de type II) lorsque tous les individus mâles peuvent être croisés avec le même échantillon d'individus femelles.

L'inconvénient de ce type d'analyse est de ne pas tirer parti de toute l'information génétique apportée par l'ensemble des individus utilisés. L'application des modèles établis en génétique quantitative y est délicate par suite de l'absence d'une partie de toutes les combinaisons possibles.

Les schémas d'analyse diallèle dans lesquels tous les croisements possibles sont réalisés fournissent plus d'information, notamment ils renseignent sur les effets réciproques. Cependant il faut remarquer que l'exécution d'un diallèle est délicat et représente un très gros travail expérimental.

Ces différents schémas d'analyse conduisent à définir deux notions importantes quel que soit le niveau de ploïdie : ce sont les notions d'aptitude générale et aptitude spécifique à la combinaison. Par contre l'expression de la variance aptitude à la combinaison en termes de variances des effets géniques n'est pas indépendante du niveau de ploïdie.

Notre but est donc avec le schéma d'analyse le plus complet, le schéma diallèle, d'arriver à un modèle permettant l'estimation des composantes de la variance génétique chez les autotétraploïdes avec le moins d'hypothèses restrictives possibles sur les relations entre allèles.

L'étude comprend deux parties, la première concerne la mise au point d'un modèle théorique, la seconde une tentative d'application sur un matériel

végétal (luzerne autotétraploïde du Puits).

La première partie sera consacrée au développement d'un modèle d'analyse génétique adapté aux autotétraploïdes à partir du modèle de décomposition statistique de la valeur génotypique déjà présenté par O.Kempthorne.

## PREMIERE PARTIE : Théorie de l'analyse diallèle chez les Tétraploïdes

### I - Définitions des croisements diallèles

On appelle "croisements diallèles complets" l'ensemble des  $p^2$  croisements possibles entre individus différents pris deux à deux, et "table diallèle" la matrice carrée  $p \times p$  dont chaque élément d'indice  $i$  et  $j$  est la mesure du résultat du croisement du  $i$ ème individu  $\phi$  par le  $j$ ème individu  $\sigma$ . La diagonale principale porte les autofécondations des  $p$  individus (fig.1).

Chacune des entrées de la table peut être soit une plante, soit un représentant d'une lignée, soit un échantillon aléatoire d'une population suivant qu'on étudie un ensemble de plantes, un ensemble de lignées ou un ensemble de populations.

Les croisements diallèles complets se répartissent en 4 groupes :

1.  $p(p-1)$  croisements et  $p$  autofécondations, c'est-à-dire la table complète.
2.  $\frac{p(p-1)}{2}$  croisements, sans les croisements réciproques, plus les  $p$  autofécondations, c'est-à-dire la moitié de la table y compris la diagonale principale.
3.  $p(p-1)$  croisements sans les autofécondations c'est-à-dire la table sans la diagonale principale.
4.  $\frac{p(p-1)}{2}$  croisements sans les croisements réciproques ni les autofécondations c'est-à-dire la moitié de la table sans la diagonale principale.

Les groupes 1 et 3 sont les plus importants car ils permettent d'estimer les effets réciproques.

$\begin{matrix} Q \\ \backslash \\ p \end{matrix}$	1	2	.....	i	.....	j	.....	p
1	$y_{11}$	$y_{12}$		$y_{1i}$		$y_{1j}$		$y_{1p}$
2	$y_{21}$	$y_{22}$		$y_{2i}$		$y_{2j}$		$y_{2p}$
⋮								
i	$y_{i1}$	$y_{i2}$		$y_{ii}$		$y_{ij}$		$y_{ip}$
⋮								
j	$y_{j1}$	$y_{j2}$		$y_{ji}$		$y_{jj}$		$y_{jp}$
⋮								
p	$y_{p1}$	$y_{p2}$		$y_{pi}$		$y_{pj}$		$y_{pp}$

Fig.1 Table diallèle

## II - ANALYSE STATISTIQUE

Deux situations peuvent se présenter dans le cas de l'étude d'un ensemble de plantes (Eisenhart 1947).

1°) les génotypes sont choisis et les conclusions tirées ne portent que sur ces génotypes. On ne pourra pas conclure sur la population, mais on pourra obtenir un classement des croisements entre eux de manière à séparer les plus intéressants pour le sélectionneur ; c'est le modèle déterministe (modèle I d'Eisenhart) appelé aussi modèle fixe.

2°) Les génotypes constituent un échantillon aléatoire de la population, cet échantillon est supposé la représenter. Les effets génotypiques sont des variables aléatoires et il est possible d'obtenir des estimations des paramètres caractérisant la structure génétique de la population : c'est le modèle aléatoire (modèle II d'Eisenhart).

Le choix du modèle à utiliser réside dans les conclusions : ou bien elles doivent se limiter aux individus effectivement étudiés (et l'expérience si elle doit être refaite devra porter sur les mêmes individus) ou bien elles doivent être étendues à la population (et l'expérience si elle doit être refaite devra porter sur de nouveaux échantillons de la population).

Dans le premier cas on choisira le modèle I, dans le second le modèle II.

Dans le modèle I les effets sont des paramètres ils peuvent donc être estimés.

Dans le modèle II ce sont des réalisations de variables aléatoires les **variances** de ces effets peuvent être estimées.

## A - Analyse de la variance "entre hybrides"

### 1°) généralités

Cette analyse est importante, chronologiquement la première, car en l'absence de différences entre les hybrides la décomposition de la variance génétique n'a pas de sens.

Le dispositif expérimental d'une étude de croisements diallèles peut comprendre  $r$  répétitions des croisements diallèles complets réalisés sur les  $p$  mêmes individus.

Le groupe (3) sera étudié ici dans le cas des blocs incomplets randomisés. Les résultats pourront être comparés à ceux du groupe (1). GRIFFING (1958) montre que lorsque les autofécondations sont incluses dans l'analyse on ne peut pas obtenir d'estimations sans biais des paramètres et variances (l'explication en sera donnée plus loin).

Les  $p(p-1)$  hybrides sont disposés aléatoirement dans chacun des  $r$  blocs, (ainsi que les  $p$  autofécondations et les  $p$  parents qui seront étudiés séparément), chaque parcelle comprend au départ un nombre constant de plantes ; par suite des accidents pouvant survenir en cours de végétation le nombre de plantes soumises aux analyses statistiques peut être variable d'une parcelle à l'autre. Cependant si la variation est faible (dans le rapport de 3 à 1) l'analyse peut être faite sur les moyennes parcellaires (Box G.E.P. 1954).

### 2°) Modèle mathématique

Soit  $y_{ijk}$  la moyenne des performances des plantes de la parcelle  $ijk$  c'est-à-dire des descendants du croisement de l'individu  $i$  par l'individu  $j$  dans le bloc  $k$ ,

le modèle que nous allons adopter est pour le groupe 3

$$y_{ijk} = \mu + v_{ij} + b_k + e_{ijk} \quad i \neq j$$

où  $\mu$  désigne une contribution commune à tous les individus

$v_{ij}$  une contribution commune à tous les descendants du croisement  $ixj$

$b_k$  une contribution commune à tous les individus du bloc  $k$

$e_{ijk}$  une déviation aléatoire

L'estimation des paramètres est faite classiquement par la méthode des moindres carrés qui consiste en la minimisation de la somme des carrés des déviations aléatoires

$$\sum_{i, j, k} e_{ijk}^2 = \sum_{\substack{ijk \\ i \neq j}} (y_{ijk} - \mu - v_{ij} - b_k)^2$$

Le système d'équations linéaire obtenu en annulant les dérivées partielles de cette somme par rapport aux différents paramètres donne les estimations de :  $\mu, v_{ij}, b_k$ . Elles n'ont de réalité que pour le modèle I mais permettent pour les deux modèles d'obtenir la décomposition de la somme des carrés des écarts à la moyenne générale :

$$\sum_{\substack{ijk \\ i \neq j}} y_{ijk}^2 - \frac{Y^2}{rp(p-1)} \quad \text{en :}$$

$$\text{somme des carrés blocs} \quad \sum_k \frac{Y^2_{\cdot \cdot k}}{p(p-1)} - \frac{Y^2}{rp(p-1)}$$

$$\text{somme des carrés croisement} \quad \sum_{ij} \frac{Y^2_{ij \cdot}}{r} - \frac{Y^2}{rp(p-1)}$$



somme des carrés résiduelle (par différence), dans lesquelles

$$Y_{...} = \sum_{\substack{ijk \\ i \neq j}} y_{ijk}, Y_{..k} = \sum_{\substack{ij \\ i \neq j}} y_{ijk}, Y_{ij.} = \sum_k y_{ijk}$$

Dans les deux modèles le calcul des espérances mathématiques a été fait ; les résultats sont réunis dans le tableau I.

A partir de cette analyse la variable F permet le test de l'hypothèse de différences entre hybrides

$$F_{\{p(p-1)-1 ; [p(p-1)-1](r-1)\}} = \frac{\text{carré moyen croisements}}{\text{carré moyen résiduel}}$$

pour le groupe (3)

$$F_{\{p^2-1 ; (r-1)(p^2-1)\}} = \frac{\text{carré moyen croisements}}{\text{carré moyen résiduel}}$$

pour le groupe (1)

## B - Etude des aptitudes à la combinaison

### 1°) généralités

Si la différence entre les croisements est significative on peut continuer la décomposition pour estimer les aptitudes générale et spécifique à la combinaison ainsi que les effets réciproques.

Une table diallèle ressemble à un plan factoriel et on pourra lui appliquer une méthode d'analyse de variance.

TABLEAU I

Variation	d.d.l.	somme des carrés	espérances mathématiques des carrés moyens	
			Modèle I	Modèle II
Blocs	r-1	$\sum_k \frac{Y^2_{..k}}{p(p-1)} - \frac{Y^2_{...}}{rp(p-1)}$	$\frac{p(p-1)}{r-1} \sum_k b_k^2 + \sigma^2$	$p(p-1)\sigma_b^2 + \sigma^2$
Hybrides	p(p-1)-1	$\sum_{ij} \frac{Y^2_{ij.}}{r} - \frac{Y^2_{...}}{rp(p-1)}$	$\frac{r}{p(p-1)-1} \sum_{i \neq j} v_{ij}^2 + \sigma^2$	$r\sigma_v^2 + \sigma^2$
Résiduelle	(r-1) x [p(p-1)-1]	par différence	$\sigma^2$	$\sigma^2$
Totale	rp(p-1)-1	$\sum_{ijk} y_{ijk}^2 - \frac{Y^2_{...}}{rp(p-1)}$	X	

a) Groupe (3)

Variation	d.d.l.	Somme des carrés	espérances mathématiques des carrés moyens	
			Modèle I	Modèle II
Blocs	r-1	$\sum_k \frac{Y^2_{..k}}{p^2} - \frac{Y^2_{...}}{rp^2}$	$\frac{p^2}{r-1} \sum_k b_k^2 + \sigma^2$	$p^2\sigma_b^2 + \sigma^2$
Hybrides	p <sup>2</sup> -1	$\sum_{ij} \frac{Y^2_{ij.}}{r} - \frac{Y^2_{...}}{rp^2}$	$\frac{r}{p^2-1} \sum_{ij} v_{ij}^2 + \sigma^2$	$r\sigma_v^2 + \sigma^2$
Résiduelle	(r-1)(p <sup>2</sup> -1)	par différence	$\sigma^2$	$\sigma^2$
Totale	rp <sup>2</sup> -1	$\sum_{ijk} y_{ijk}^2 - \frac{Y^2_{...}}{rp^2}$	X	

b) Groupe (1)

L'analyse de la variance selon le schéma de Yates (1947) est valable quels que soient les parents étudiés, leur homozygotie, leur polyploidie elle est commune aux deux modèles.

Cette analyse permet l'étude des aptitudes à la combinaison qui est un concept important pour la sélection.

Les définitions de l'aptitude générale à la combinaison et de l'aptitude spécifique à la combinaison ont été données par SPRAGUE et TATUM (1942).

L'aptitude générale à la combinaison désigne la performance moyenne d'une lignée (ou d'une plante) en combinaison avec l'ensemble des autres lignées (ou plantes).

L'aptitude spécifique à la combinaison désigne la performance particulière d'un croisement inférieur ou supérieur à la performance moyenne attendue sur la base de l'aptitude générale.

GRIFFING (1956) donne pour chaque groupe de croisements diallèles et pour chaque modèle le tableau d'analyse de la variance. Le terme  $v_{ij}$  peut être décomposé en vue de mettre en évidence les aptitudes à la combinaison :

$$v_{ij} = g_i + g_j + s_{ij} + r_{ij}$$

$g_i$  désigne la contribution apportée par le parent  $i$

$s_{ij}$  une contribution spécifique au croisement de  $i$  par  $j$   $s_{ij} = s_{ji}$

$r_{ij}$  une contribution provenant de la différence entre le croisement et son croisement réciproque  $r_{ij} = -r_{ji}$

Le modèle envisagé peut donc s'écrire :

$$y_{ijk} = \mu + g_i + g_j + s_{ij} + r_{ij} + b_k + e_{ijk}$$

dans le modèle I  $g_i$ ,  $s_{ij}$ ,  $r_{ij}$ ,  $b_k$  sont supposés être des paramètres certains le matériel expérimental est considéré comme la population sur laquelle les

inductions sont faites. Le but en est la comparaison des aptitudes à la combinaison des parents quand les parents eux-mêmes sont utilisés comme testeurs, et l'identification des croisements de meilleurs rendements. Ainsi nous sommes particulièrement intéressés par l'estimation des effets "aptitude à la combinaison" et par le calcul approprié des écarts types des différences entre les effets.

Dans le modèle II  $g_i, s_{ij}, r_{ij}, b_k$  sont des variables aléatoires l'hypothèse est que nous avons affaire à des échantillons aléatoires d'une certaine population parente et les inductions ne doivent pas être faites sur les lignées individuelles dans l'échantillon, mais sur les paramètres de la population parente. En particulier nous sommes intéressés par l'estimation des composantes génétique et d'environnement de la variance de la population complexe. Pour cela nous supposons que les effets dans ce modèle sauf  $\mu$  sont normalement et indépendamment distribués de moyennes nulles et de variances  $\sigma_\theta^2$  ( $\theta = b, g, s, r$ ). Les estimations des composantes de la variance sont alors obtenues en égalant les carrés moyens observés et attendus dans l'analyse de variance. Les écarts types des estimations des composantes de la variance sont calculés à partir des variances des carrés moyens.

Le calcul des sommes des carrés des effets est conduit par la méthode des moindres carrés, sans faire d'hypothèse sur le sens des paramètres.

On obtient :

$$\text{sommes des carrés "blocs"} S_b = \frac{\sum_k Y_{..k}^2}{p(r-1)} - \frac{Y_{...}^2}{rp(p-1)}$$

somme des carrés "aptitude générale à la combinaison"

$$S_g = \frac{\sum_i (Y_{i..} + Y_{.i.})^2}{2r(p-2)} - \frac{2Y_{...}^2}{rp(p-2)}$$

somme des carrés "aptitude spécifique à la combinaison"

$$S_s = \sum_{i < j} \frac{(Y_{ij.} + Y_{ji.})^2}{2r} - \sum_i \frac{(Y_{i..} + Y_{.i.})^2}{2r(p-2)} + \frac{Y_{...}^2}{r(p-1)(p-2)}$$

somme des carrés "effets réciproques"

$$S_r = \sum_{i < j} \frac{(Y_{ij.} - Y_{ji.})^2}{2r}$$

Nous allons examiner ces sommes des carrés dans chacun des deux modèles.

Modèle I :

Les estimations des effets sont les suivantes (annexe 1)

$$\begin{aligned} \hat{\mu} &= \frac{Y_{...}}{rp(p-1)} & \hat{\alpha}_i &= \frac{Y_{i..} + Y_{.i.}}{2r(p-2)} - \frac{Y_{...}}{rp(p-2)} \\ \hat{\beta}_k &= \frac{Y_{..k}}{p(p-1)} - \frac{Y_{...}}{rp(p-1)} & \hat{s}_{ij} &= \frac{Y_{ij.} + Y_{ji.}}{2r} - \frac{Y_{i..} + Y_{.i.} + Y_{j..} + Y_{.j.}}{2r(p-2)} + \\ & & & \frac{Y_{...}}{r(p-1)(p-2)} \\ \hat{r}_{ij} &= \frac{Y_{ij.} - Y_{ji.}}{2r} \end{aligned}$$

Les espérances mathématiques des carrés moyens se calculent selon une méthode classique (annexe 2), on obtient le tableau II.

Pour le groupe (1) seul le tableau final est donné.

TABLEAU II

variation	d.d.l.	Somme des carrés	Espérances mathématiques des carrés moyens
Blocs	$r-1$	$\sum_k \frac{y_{\dots k}^2}{p(p-1)} - \frac{y_{\dots \dots}^2}{rp(p-1)}$	$\sigma^2 + \frac{p(p-1)}{r-1} \sum_k b_k^2$
AGC	$p-1$	$\sum_i \frac{(y_{i \dots} + y_{\dots i})^2}{2r(p-2)} - \frac{2y_{\dots \dots}^2}{rp(p-2)}$	$\sigma^2 + \frac{2r(p-2)}{p-1} \sum_i g_i^2$
ASC	$\frac{p(p-3)}{2}$	$\sum_{i < j} \frac{(y_{ij \dots} + y_{ji \dots})^2}{2r} - \frac{\sum (y_{i \dots} + y_{\dots i})^2}{2r(p-2)} + \frac{y_{\dots \dots}^2}{r(p-1)(p-2)}$	$\sigma^2 + 2r \frac{2}{p(p-3)} \sum_{i < j} s_{ij}^2$
Réciproques	$\frac{p(p-1)}{2}$	$\sum_{i < j} \frac{(y_{ij \dots} - y_{ji \dots})^2}{2r}$	$\sigma^2 + 2r \frac{2}{p(p-1)} \sum_{i < j} r_{ij}^2$
Résiduelle	$v = (r-1)p(p-1)-1$	par différence	$\sigma^2$
Totale	$rp(p-1) - 1$	$\sum_{ijk} y_{ijk}^2 - \frac{y_{\dots \dots}^2}{rp(p-1)}$	

a) Groupe (3)

TABLEAU II (suite)

Variation	d.d.l.	Somme des carrés	Espérances mathématiques des carrés moyens
Blocs	$r-1$	$\sum_k \frac{Y_{..k}^2}{p^2} - \frac{Y_{...}^2}{p^2 r}$	$\sigma^2 + \frac{p^2}{r-1} \sum_k b_k^2$
A G C	$p-1$	$\sum_i \frac{(Y_{i..} + Y_{.i.})^2}{2pr} - \frac{2Y_{...}^2}{p^2 r}$	$\sigma^2 + \frac{2pr}{p-1} \sum_i g_i^2$
A S C	$\frac{p(p-1)}{2}$	$\sum_{i,j} \frac{Y_{ij.}(Y_{ij.} + Y_{ji.})}{2r} + \sum_i \frac{(Y_{i..} + Y_{.i.})^2}{2pr} + \frac{Y_{...}^2}{p^2 r}$	$\sigma^2 + \frac{2pr}{p(p-1)} \sum_{i,j} s_{ij}^2$
Réciproques	$\frac{p(p-1)}{2}$	$\sum_{i < j} \frac{(Y_{ij.} - Y_{ji.})^2}{2r}$	$\sigma^2 + 2r \frac{2}{p(p-1)} \sum_{i < j} r_{ij}^2$
Résiduelle	$v = (r-1)(p^2-1)$	par différence	$\sigma^2$
Totale	$rp^2-1$	$\sum_{ijk} y_{ijk}^2 - \frac{Y_{...}^2}{p^2 r}$	

b) Groupe (1)

Les tests F de nullité des différents effets sont les suivants :

$$\Lambda.G.C. \quad F_{[p-1], v} = \frac{\text{Carré moyen } \Lambda.G.C.}{\text{Carré moyen résiduel}}$$

$$\Lambda.S.C. \quad F_{\left[\frac{p(p-3)}{2}, v\right]} = \frac{\text{Carré moyen } \Lambda.S.C.}{\text{Carré moyen résiduel}}$$

$$(F_{\left[\frac{p(p-1)}{2}, v\right]} = \frac{\text{Carré moyen } \Lambda.S.C.}{\text{Carré moyen résiduel}} \quad \text{pour le groupe (1)})$$

$$\text{Réciproque } F_{\left[\frac{p(p-1)}{2}, v\right]} = \frac{\text{Carré moyen réciproque}}{\text{Carré moyen résiduel}}$$

Dans le modèle I il est intéressant de connaître les variances des estimations des paramètres et de certaines différences entre ces paramètres de façon à connaître la précision de ces estimations et à classer les génotypes suivant leurs aptitudes (annexe 3).

Pour le groupe (3)  $\text{var}(y_{ijk}) = \sigma^2 = \text{carré moyen résiduel}$

$$\text{var}(y_{ijk} - y_{lmn}) = 2\sigma^2$$

$$\text{var}(\hat{\mu}) = \frac{\sigma^2}{rp(p-1)}$$

$$\text{var}(\hat{g}_i) = \frac{p-1}{2rp(p-2)} \sigma^2$$

$$\text{var}(\hat{s}_{ij}) = \frac{p-3}{2r(p-1)} \sigma^2$$

$$i \neq j$$

$$\text{var}(\hat{g}_i - \hat{g}_j) = \frac{\sigma^2}{r(p-2)}$$

$$i \neq j$$

$$\text{var}(\hat{s}_{ij} - \hat{s}_{ik}) = \frac{p-3}{r(p-2)} \sigma^2$$

$$i \neq j, k ; j \neq k$$



$$\text{var } (\hat{s}_{ij} - \hat{s}_{kl}) = \frac{p-4}{r(p-2)} \hat{\sigma}^2$$

$$\begin{aligned} i &\neq j, k, l, \\ j &\neq k, l ; k \neq l \end{aligned}$$

$$\text{var } (\hat{r}_{ij}) = \frac{\hat{\sigma}^2}{2r}$$

pour le groupe (1)

$$\text{var } (y_{jk}) = \hat{\sigma}^2 = \text{carré moyen résiduel}$$

$$\text{var } (y_{ijk} - y_{lmn}) = 2 \hat{\sigma}^2$$

$$\text{var } (\hat{\mu}) = \frac{2\hat{\sigma}^2}{p^2 r}$$

$$\text{var } (\hat{g}_i) = \frac{p-1}{2p^2 r} \hat{\sigma}^2$$

$$\text{var } (\hat{s}_{ii}) = \frac{(p-1)^2}{p^2 r} \hat{\sigma}^2$$

$$\text{var } (\hat{s}_{ij}) = \frac{p^2 - 2p + 2}{2rp^2} \hat{\sigma}^2$$

$$i \neq j$$

$$\text{var } (\hat{g}_i - \hat{g}_j) = \frac{\hat{\sigma}^2}{pr}$$

$$i \neq j$$

$$\text{var } (\hat{s}_{ii} - \hat{s}_{jj}) = \frac{2(p-2)}{pr} \hat{\sigma}^2$$

$$i \neq j$$

$$\text{var } (\hat{s}_{ii} - \hat{s}_{ij}) = \frac{3p-2}{2pr} \hat{\sigma}^2$$

$$i \neq j$$

$$\text{var } (\hat{s}_{ij} - \hat{s}_{jk}) = \frac{3(p-2)}{2pr} \hat{\sigma}^2$$

$$i \neq j, k ; j \neq k$$

$$\text{var } (\hat{s}_{ij} - \hat{s}_{ik}) = \frac{p-1}{pr} \hat{\sigma}^2$$

$$i \neq j, k ; j \neq k$$

$$\text{var } (\hat{s}_{ij} - \hat{s}_{kl}) = \frac{p-2}{pr} \hat{\sigma}^2$$

$$i \neq j, k, l ; j \neq k ; k \neq l$$

$$\text{var } (\hat{r}_{ij}) = \frac{\hat{\sigma}^2}{2r} \quad i \neq j$$

### Modèle II :

Dans ce modèle les valeurs observées sont des réalisations de variables aléatoires et n'ont en tant que telles que peu de sens ; elles ne sont pas des quantités estimables (O. KEMP THORNE 1952) cependant la décomposition de la somme des carrés reste valable.

Le calcul des espérances mathématiques des carrés moyens des effets est moins simple que dans le modèle I (annexe 4) on obtient le tableau III pour le modèle II

Les paramètres qui nous intéressent dans ce modèle sont les variances  $\sigma_g^2$ ,  $\sigma_s^2$ ,  $\sigma_r^2$ . Pour tester leur nullité on utilise la variable F.

#### Pour le groupe (3)

test de $\sigma_g^2 = 0$	$F \left[ (p-1), \frac{p(p-3)}{2} \right]$	$\frac{\text{carré moyen "AGC"}}{\text{carré moyen "ASC"}}$
test de $\sigma_s^2 = 0$	$F \left[ \frac{p(p-3)}{2}, v \right]$	$\frac{\text{carré moyen "ASC"}}{\text{carré moyen résiduel}}$
test de $\sigma_r^2 = 0$	$F \left[ \frac{p(p-1)}{2}, v \right]$	$\frac{\text{carré moyen "réciproque"}}{\text{carré moyen résiduel}}$

Les estimations des composantes de la variance sont les suivantes :

$$\hat{\sigma}_g^2 = \frac{1}{2r(p-2)} (CM_g - CM_s)$$

$$\hat{\sigma}_s^2 = \frac{1}{2r} (CM_s - CM_{rés})$$

$$\hat{\sigma}_r^2 = \frac{1}{2r} (CM_r - CM_{rés})$$

TABLEAU III

variation	d.d.l.	Somme des carrés	Espérances mathématiques des carrés moyens
Blocs	r-1	$\sum_k \frac{Y_{\dots k}^2}{p(p-1)} - \frac{Y_{\dots \dots}^2}{rp(p-1)}$	$\sigma^2 + p(p-1)\sigma_b^2$
A G C	p-1	$\sum_i \frac{(Y_{i..} + Y_{.i.})^2}{2r(p-2)} - \frac{2Y_{\dots \dots}^2}{rp(p-2)}$	$\sigma^2 + 2r\sigma_s^2 + 2r(p-2)\sigma_g^2$
A S C	$\frac{p(p-3)}{2}$	$\sum_{i < j} \frac{(Y_{ij.} + Y_{ji.})^2}{2r} - \sum_i \frac{(Y_{i..} + Y_{.i.})^2}{2r(p-2)} + \frac{Y_{\dots \dots}^2}{r(p-1)(p-2)}$	$\sigma^2 + 2r\sigma_s^2$
Réciproq.	$\frac{p(p-1)}{2}$	$\sum_{i < j} \frac{(Y_{ij.} - Y_{ji.})^2}{2r}$	$\sigma^2 + 2r\sigma_r^2$
Résidu	$v = (r-1)(p(p-1)-1)$	par différence	$\sigma^2$
Totale	$rp(p-1) - 1$	$\sum_{ijk} y_{ijk}^2 - \frac{Y_{\dots \dots}^2}{rp(p-1)}$	

Groupe (3)

on peut donner une approximation de la variance de l'estimation d'une variance par la formule  $\text{var} (CM_{\theta}) = \frac{2}{v_{\theta}} (CM_{\theta})^2$

(SCHEFFE 1961)

$v_{\theta}$  étant le nombre de degré de liberté du carré moyen  $CM_{\theta}$ , on a donc

$$\text{var}(\hat{\sigma}_g^2) = \frac{1}{4r^2(p-2)^2} \left[ \frac{2}{p-1} (CM_g)^2 + \frac{4}{p(p-3)} (CM_s)^2 \right]$$

$$\text{var}(\hat{\sigma}_s^2) = \frac{1}{4r^2} \left[ \frac{4}{p(p-3)} (CM_s)^2 + \frac{2}{v} (CM_{rés})^2 \right]$$

$$\text{var}(\hat{\sigma}_r^2) = \frac{1}{4r^2} \left[ \frac{4}{p(p-1)} (CM_r)^2 + \frac{2}{v} (CM_{rés})^2 \right] \quad (\text{annexe 5})$$

On peut donner aussi un intervalle de confiance approximatif de l'estimation d'une variance (Annexe 11)

Pour le Groupe (I) GRIFFING (1958) montre que les estimations des aptitudes générale et spécifique à la combinaison ne peuvent être faites.

En effet si on peut prendre le modèle

$$y_{ijk} = \mu + g_i + g_j + s_{ij} + r_{ij} + b_k + e_{ijk}$$

pour les hybrides il n'en est pas de même pour les individus issus d'autofécondations. Pour ces derniers le modèle adopté devrait être :

$$Y_{iik} = \mu_I + h_{ii} + b_k + e_{iik}$$

$$E(h_{ii}) = \sigma_I^2 \quad E(h_{ii}g_i) = \sigma_{gI}$$

avec ce modèle les carrés moyens "aptitude générale" et "aptitude spécifique" ont pour espérance mathématique :

$$E (CM_g) = \sigma^2 + r \left[ \frac{2}{p} \sigma_I^2 + \frac{2}{p} (p-2)^2 \sigma_g^2 + \frac{2}{p} (p-2) \sigma_s^2 + \frac{4(p-2)}{p} \sigma_{gI} \right]$$

$$E (CM_s) = \sigma^2 + r \left[ \frac{2}{p} (\mu_I - \mu)^2 + \frac{2}{p^2} (p-1) \sigma_I^2 + \frac{8}{pr} (p-1) \sigma_g^2 + \frac{2}{p^2} (p^2 - 2p + 2) \sigma_s^2 + \frac{8}{p^2} (p-1) \sigma_{gI} \right]$$

Les expressions sont beaucoup plus compliquées, et les nouveaux paramètres introduits  $\mu_I$ ,  $\sigma_I^2$ ,  $\sigma_{gI}$  ne peuvent pas être estimés directement par cette analyse de variance.

Dès cette étape il est possible de savoir s'il y a des différences génétiques significatives entre les croisements réalisés (c'est-à-dire si les aptitudes à la combinaison sont significativement différentes entre plantes) si ces différences existent avec le modèle I il est possible de classer les croisements entre eux selon leur moyenne, leur aptitude spécifique ou leur aptitude générale à la combinaison pour les différents caractères étudiés.

Dans le modèle II il sera possible par l'utilisation d'un modèle génétique approprié d'étudier la structure génétique de la population dont sont issus tous les parents étudiés pour le ou les caractères considérés.

### C - Décomposition des effets réciproques

Seul le groupe (3) sera envisagé ici. Si les effets réciproques existent on peut mettre en évidence la part de ces effets due au parent femelle qu'on appelle effet maternel (COCKERHAM 1963).

Le paramètre de l'effet réciproque  $r_{ij}$  peut être décomposé de la façon suivante :

$$r_{ij} = m_i - m_j + r'_{ij}$$

$m_i$  est une variable aléatoire de moyenne nulle et de variance  $\sigma_m^2$

représentant un effet dû à l'individu  $i$  en tant que mère ou effet maternel ;

$r'_{ij}$  est une variable aléatoire de moyenne nulle et de variance  $\sigma_r^2$ ,

effet réciproque autre que l'effet maternel.

La décomposition est la suivante :

TABLEAU IV

Variation	d.d.l.	Somme des carrés	Espérances mathématiques des carrés moyens
Effet maternel	$p-1$	$\sum_i \frac{(Y_{i..} - Y_{..i})^2}{2rp}$	$\sigma^2 + r\sigma_r^2 + rp\sigma_m^2$
Effet réciproque ( $r'$ )	$(p-1)(p-2)/2$	par différence	$\sigma^2 + r\sigma_r^2$
Effet réciproque ( $r$ )	$p(p-1)/2$	$\sum_{i < j} \frac{(Y_{ij.} - Y_{ji.})^2}{2r}$	

Les tests de nullité de  $\sigma_r^2$ , et de  $\sigma_m^2$  sont les suivants :

$$\sigma_r^2 = 0 \quad F \left[ \frac{(p-1)(p-2)}{2} ; v \right] = \frac{\text{carré moyen effet réciproque } (r')}{\text{carré moyen résiduel}}$$

$$\sigma_m^2 = 0 \quad F \left[ (p-1) ; \frac{(p-1)(p-2)}{2} \right] = \frac{\text{carré moyen effet maternel}}{\text{carré moyen effet réciproque}}$$

dans le cas où  $\sigma_r^2$ , et  $\sigma_m^2$  sont significativement différents de zéro les aptitudes générale et spécifique devront être diminuées respectivement de  $\sigma_m^2/2$  et  $\sigma_r^2$ ,

TABLEAU IV (bis)

Modèle :  $y_{ijk} = \mu + g_i^v + g_j^v + s_{ij}^v + m_i - m_j + r_{ij}^v + b_k + e_{ijk}$

$s_{ij}^v = s_{ji}^v$ ,  $g^v$   $s^v$   $m$   $r^v$   $b$   $e$  sont des variables aléatoires de moyenne nulle et de variance  $\sigma_g^2$   $\sigma_s^2$   $\sigma_m^2$   $\sigma_r^2$   $\sigma_b^2$   $\sigma^2$

Variation	d.d.l.	Espérances des carrés moyens
Blocs	$r-1$	
A. G. C.	$p-1$	$\sigma^2 + r\sigma_r^2 + 2r\sigma_s^2 + r(p-2)\sigma_m^2 + 2r(p-2)\sigma_g^2$
A. S. C.	$p(p-3)/2$	$\sigma^2 + r\sigma_r^2 + 2r\sigma_s^2$
maternelle	$p-1$	$\sigma^2 + r\sigma_r^2 + rp\sigma_m^2$
résiduelle réciproque	$p(p-2)/2$	$\sigma^2 + r\sigma_r^2$
résiduelle	$(r-1)(p^2-p-1)$	$\sigma^2$

Pour éprouver l'hypothèse  $\sigma_g^2 = 0$  on peut (SATTERTHWAITE 1946) faire un test approximatif, le nombre de degrés de liberté à attribuer à une combinaison

linéaire de carrés moyens  $\sum_i a_i CM_i$  est  $f' = \frac{(\sum_i a_i CM_i)^2 / \sum_i \frac{a_i^2 CM_i^2}{f_i}}{f_i}$

$f_i$  étant le nombre de degrés de liberté de  $CM_i$

D - Etude de la variance intra-parcellaire

Nous avons jusqu'à présent utilisé les moyennes des parcelles mais chaque parcelle contient plusieurs plantes et il est intéressant pour la suite d'obtenir une estimation de la variance intra-parcellaire d'où la nécessité de faire une analyse de variance entre et intra-parcelles, soit :

$x_{ijk1}$  le résultat de la mesure faite sur un individu de la parcelle "ijk" ; et  $n_{ijk}$  le nombre d'individus de cette parcelle ; le  $y_{ijk}$  utilisé plus haut est donc  $\sum_1 \frac{x_{ijk1}}{n_{ijk}}$

$$\text{soit } X_{ijk.} = \sum_1 x_{ijk1} \quad X_{....} = \sum_{ijk1} x_{ijk1} \quad N_{...} = \sum_{ijk} n_{ijk}$$

Le tableau de cette analyse de variance est le suivant :

TABEAU V

source de variation	d.d.l.	somme des carrés	carré moyen
entre parcelles	$rp(p-1)-1$	$\sum_{ijk} \frac{x_{ijk.}^2}{n_{ijk}} - \frac{X_{....}^2}{N_{...}}$	
intra parcelles	$N_{...} - rp(p-1)$	par différence	$CM_I$
T o t a l	$N_{...} - 1$	$\sum_{ijk1} x_{ijk1}^2 - \frac{X_{....}^2}{N_{...}}$	

cette décomposition est intéressante puisque l'espérance mathématique du carré moyen de la variance intra-parcellaire s'exprime en fonction de la covariance entre frères, chaque parcelle contenant des individus frères.



# E - Analyse des autofécondations

Soit  $x_{ikl}$  la performance du lième descendant issu de l'autofécondation de l'individu i dans le k ième bloc.

$$\text{Soit } y_{ik} = \sum_1 \frac{x_{ikl}}{n_{ik}}$$

$y_{ik}$  est la moyenne des  $n_{ik}$  descendants issus de l'autofécondation de l'individu i dans le k ième bloc.

Le modèle mathématique est le suivant :

$$y_{ik} = \mu_I + h_i + b_k + e_{ik}$$

$h_i$  contribution commune à tous les descendants issus de l'autofécondation de i

$b_k$  contribution commune à tout le bloc k

$e_{ik}$  déviation aléatoire

$h_i, b_k, e_{ik}$  sont des variables aléatoires indépendantes de moyenne nulle et variance  $\sigma_h^2, \sigma_b^2, \sigma^2$ .

$$x_{ikl} = \mu_I + G_{il} + b_k + f_{ikl} + e_{ik}^*$$

$$E(G_{il}) = 0, \quad E(G_{il}^2) = \sigma_I^2, \quad E(G_{il} G_{il'}) = \text{cov}(\text{FS})_I$$

$f_{ikl}$  effet d'environnement propre à l'individu ikl

$e_{ik}^*$  effet d'environnement propre à la parcelle ik

Une analyse de variance blocs, descendants, résidu, donne une estimation de  $\sigma_h^2$ .

Une analyse de variance "entre et intra-parcelles" donne une estimation de  $\sigma_I^2 + \sigma_f^2 - \text{cov}(\text{FS})_I$  (carré moyen intraparcellaire).

On a la relation suivante :

$$\frac{\hat{\sigma}_I^2 - \overbrace{\text{Cov}(\text{FS})_I}}{\quad} + \overbrace{\text{Cov}(\text{FS})_I} = \hat{\sigma}_h^2$$

n étant le nombre moyen (harmonique) d'individus par parcelle.

Finalement :

$$\widehat{\text{Cov}}(\text{FS})_{\text{I}} - \frac{\hat{\sigma}_{\text{f}}^2}{n} = \hat{\sigma}_{\text{h}}^2 - \frac{\text{CM}_{\text{I}}^{\text{K}}}{n}$$

$$\hat{\sigma}_{\text{I}}^2 + \frac{n-1}{n} \hat{\sigma}_{\text{f}}^2 = \frac{n-1}{n} \text{CM}_{\text{I}}^{\text{K}} - \hat{\sigma}_{\text{h}}^2$$

### III - Analyse génétique

#### A - modèle génétique

Nous allons nous placer dans le cas du modèle II (groupe (3)) qui est le plus intéressant pour les considérations génétiques :

Les hypothèses sont les suivantes :

- 1 - ségrégation tétraploïde chromosomique ( $\alpha = 0$ ) (1)
- 2 - épistasie quelconque et absence de linkage.
- 3 - nombre d'allèles quelconque à chaque locus.
- 4 - parents consanguins ou non, issus d'une population panmictique sans sélection à l'équilibre.
- 5 - pas d'interaction génotype x environnement.

La structure génétique des données d'un diallèle est :

- 1°) les individus à l'intérieur d'une parcelle sont frères.
- 2°) les individus de répétitions différentes d'un croisement particulier sont frères.
- 3°) les individus d'une même répétition ou de deux répétitions différentes ayant un parent en commun sont demi-frères.

#### B - Calcul des espérances mathématiques des carrés moyens en fonction du modèle génétique.

Pour interpréter l'analyse de la variance nous devons trouver les espérances des carrés moyens en fonction de notre modèle génétique. Les parents sont un échantillon aléatoire d'un ensemble d'individus issus d'une population panmictique par autofécondations jusqu'à un certain point mesuré par le coefficient de consanguinité  $F$ . On doit noter que le cas où  $F = 0$  correspond à la

---

(1)  $\alpha$  est l'indice de séparation, c'est-à-dire la proportion de division pseudo-équationnelle à un locus donné. (Y. DEIARLY 1963).

situation où les parents sont choisis dans une population panmictique sans consanguinité et que  $F = 1$  correspond au cas où les parents sont des individus homozygotes résultant par autofécondations jusqu'à l'homozygotie d'une population panmictique. Il apparaît que la continuité totale des possibilités depuis  $F = 0$  jusqu'à  $F = 1$  est intéressante.

La notation utilisée est celle des auteurs anglo-saxons :

covariance entre frères  $= \text{cov}(\text{FS})$

covariance entre demi-frères  $= \text{cov}(\text{HS})$

covariance entre parent et enfant  $= \text{cov}(\text{PO})$

O. KEMPTHORNE (1957) utilise une décomposition de la déviation aléatoire attachée à chaque individu en une déviation  $e_{ijk}$  commune à tous les individus d'une parcelle et une déviation  $f_{ijkl}$  au niveau de l'individu.

On peut écrire :

$$x_{ijkl} = \mu + G_{ij1} + b_k + e_{ijk} + f_{ijkl}$$

$b_k$ ,  $e_{ijk}$ ,  $f_{ijkl}$  sont des variables aléatoires indépendantes normalement distribuées de moyenne nulle et de variance  $\sigma_b^2$ ,  $\sigma_e^2$ ,  $\sigma_f^2$ .

$G_{ij1}$  est un effet global génotypique l'indice 1 est nécessaire puisque deux frères n'ont pas des génotypes totalement identiques.

On montre que (annexe 6 et 7) l'espérance mathématique du carré moyen intra-parcellaire peut se mettre sous la forme

$$\sigma_f^2 + \sigma_G^2 - \text{cov}(\text{FS})$$

Nous avons supposé que la variance génétique d'un descendant est  $\sigma_G^2$  variance génétique de la population originale, ceci est vrai sous les hypothèses d'absence de linkage et de sélection.

La variance de la moyenne de la parcelle "ijk" est (annexe 8)

$$\sigma_b^2 + \sigma_2^2 + \frac{\sigma_1^2}{n_{ijk}} \quad \text{avec} \quad \sigma_1^2 = \sigma_f^2 + \sigma_G^2 - \text{cov}(\text{FS}) \quad \text{et}$$

$$\sigma_2^2 = \sigma_e^2 + \text{cov}(\text{FS})$$

La covariance entre deux moyennes de descendants d'un même croisement dans des répétitions différentes est  $\text{cov}(\text{FS})$ .

La covariance entre deux moyennes de descendants ayant un parent commun dans des répétitions différentes est  $\text{cov}(\text{HS})$ .

A l'aide de ces données on obtient les espérances mathématiques des différents carrés moyens de l'analyse de variance en fonction des covariances entre frères et entre demi-frère (annexe 9) d'où le tableau VI

Les estimations des covariances sont donc :

$$\widehat{\text{cov}}(\text{HS}) = \sigma_g^2,$$

$$\widehat{\text{cov}}(\text{FS}) = 2\hat{\sigma}_g^2 + \hat{\sigma}_s^2 + \hat{\sigma}_r^2/2 \quad \text{et}$$

$$\hat{\sigma}_G^2 + \hat{\sigma}_f^2 = \text{CM}_I + \widehat{\text{cov}}(\text{FS})$$

TABLEAU VI

Variation	C.M.	Espérance stat. (sur les moyennes parcellaires)	Espérance génétique
Blocs	$CM_b$	$\sigma^2 + p(p-1)\sigma_b^2$	$\sigma_2^2 - covFS + n_h \sigma_1^2 + p(p-1)\sigma_b^2$
A.G.C.	$CM_g$	$\sigma^2 + 2r\sigma_s^2 + 2r(p-2)\sigma_g^2$	$\sigma_2^2 - covFS + n_h \sigma_1^2 + 2rcovFS + 2r(p-4)covHS - r\sigma_r^2$
A.S.C.	$CM_s$	$\sigma^2 + 2r\sigma_s^2$	$\sigma_2^2 - covFS + n_h \sigma_1^2 + 2rcovFS - 4rcovHS - r\sigma_r^2$
Effet réc.	$CM_r$	$\sigma^2 + 2r\sigma_r^2$	$\sigma_2^2 - covFS + n_h \sigma_1^2 + 2r\sigma_r^2$
Résiduelle	$CM_{rés}$	$\sigma^2$	$\sigma_2^2 - covFS + n_h \sigma_1^2$

C - Définitions des variances d'additivité, de dominance, d'épistasie

Nous allons maintenant décomposer la variance génétique pour mettre en évidence les variances d'additivité, de dominance et d'épistasie dans le cas de la tétraploïdie.

1°) cas d'un locus

a) Définitions statistiques des effets géniques

Le génotype d'un individu tétraploïde s'écrit :

$$A_i A_j A_k A_l$$

$A_i$  désignant un allèle

La population est en équilibre en panmixie et la fréquence du gène  $A_i$  est  $p_i$ .

La composition génétique de la population peut s'écrire symboliquement

$$\left(\sum_i p_i A_i\right)^4 = \sum_{ijkl} p_i p_j p_k p_l A_i A_j A_k A_l$$

Chaque gène  $A_i$  apporte une contribution à la mesure du caractère qu'on étudie, on peut donc remplacer chaque gène par une valeur et par conséquent chaque génotype aussi, l'expression précédente devient alors algébrique et représente la moyenne de la population panmictique.

$$\mu = \rho^4 = \sum_{ijkl} p_i p_j p_k p_l A_i A_j A_k A_l$$

$$\text{avec } \rho = \sum_i p_i A_i$$

$$\text{posons } u_i = A_i - \rho$$

$u_i$  représente l'écart de l'effet du gène  $A_i$  à la moyenne  $\rho$  des effets des allèles.  $u_i$  est donc composé d'un effet moyen commun à tous les allèles et d'un effet particulier au gène  $A_i$ .

$$\text{On peut écrire : } A_i A_j A_k A_l = (u_i + \rho) (u_j + \rho) (u_k + \rho) (u_l + \rho)$$

ou en développant le produit :

$$\begin{aligned} \Lambda_i \Lambda_j \Lambda_k \Lambda_l &= \rho^4 + \rho^3 u_i + \rho^3 u_j + \rho^3 u_k + \rho^3 u_l + \rho^2 u_i u_j + \rho^2 u_i u_k + \rho^2 u_i u_l + \rho^2 u_j u_k \\ &\quad + \rho^2 u_j u_l + \rho^2 u_k u_l + \rho u_i u_j u_k + \rho u_i u_j u_l + \rho u_i u_k u_l + \rho u_j u_k u_l \\ &\quad + u_i u_j u_k u_l \end{aligned}$$

1°) Posons  $\alpha_i = \rho^3 u_i$

L'ensemble des génotypes contenant  $\Lambda_i$  peut s'écrire symboliquement :

$$\Lambda_i \left( \sum_s p_s \Lambda_s \right)^3$$

algébriquement cette expression est la moyenne des valeurs des génotypes contenant  $\Lambda_i$ , si on centre par rapport à la moyenne générale on obtient :

$$\begin{aligned} \Lambda_i \left( \sum_s p_s \Lambda_s \right)^3 - \mu &= \\ \Lambda_i \left( \sum_s p_s \Lambda_s \right)^3 - \left( \sum_s p_s \Lambda_s \right)^4 &= \\ \left( \sum_s p_s \Lambda_s \right)^3 (\Lambda_i - \sum_s p_s \Lambda_s) &= \rho^3 u_i \end{aligned}$$

on voit donc que  $\alpha_i$  représente la moyenne centrée des performances de tous les génotypes contenant  $\Lambda_i$ .

$\alpha_i$  est l'effet additif du gène  $\Lambda_i$  ou effet principal par analogie à l'analyse des plans factoriels.

2°) Posons  $\beta_{ij} = \rho^2 u_i u_j$

L'ensemble des génotypes contenant l'association de gènes  $\Lambda_i \Lambda_j$  peut s'écrire symboliquement :

$$\Lambda_i \Lambda_j \left( \sum_s p_s \Lambda_s \right)^2$$

algébriquement cette expression est la moyenne des valeurs des génotypes contenant l'association  $\Lambda_i \Lambda_j$ . Pour obtenir l'effet spécifique à l'association  $\Lambda_i \Lambda_j$



(tout comme en analyse des plans factoriels) il faut retrancher de cette expression la moyenne générale et les effets principaux.

$$\text{Soit } A_i A_j \left( \sum_s p_s A_s \right)^2 - \alpha_i - \alpha_j - \mu =$$

$$A_i A_j \left( \sum_s p_s A_s \right)^2 - (A_i - \rho) \left( \sum_s p_s A_s \right)^3 - (A_j - \rho) \left( \sum_s p_s A_s \right)^3 - \left( \sum_s p_s A_s \right)^4 =$$

$$\left( \sum_s p_s A_s \right)^2 \left[ A_i A_j - \left( \sum_s p_s A_s \right) A_i + \left( \sum_s p_s A_s \right)^2 - \left( \sum_s p_s A_s \right) A_j + \left( \sum_s p_s A_s \right)^2 - \left( \sum_s p_s A_s \right)^2 \right] =$$

$$\left( \sum_s p_s A_s \right)^2 \left[ A_i (A_j - \sum_s p_s A_s) - \left( \sum_s p_s A_s \right) (A_j - \sum_s p_s A_s) \right] =$$

$$\left( \sum_s p_s A_s \right)^2 (A_i - \sum_s p_s A_s) (A_j - \sum_s p_s A_s) = \rho^2 u_i u_j$$

On voit que  $\beta_{ij}$  représente un écart propre à l'association  $A_i A_j$

$\beta_{ij}$  est l'effet digénique spécifique associé aux gènes  $A_i$  et  $A_j$  c'est l'équivalent d'une interaction du premier ordre.

$$3^\circ) \text{ Posons } \gamma_{ijk} = \rho u_i u_j u_k$$

Ici encore un raisonnement analogue peut s'appliquer. L'ensemble des génotypes contenant l'association  $A_i A_j A_k$  peut s'écrire symboliquement :

$$A_i A_j A_k \left( \sum_s p_s A_s \right)$$

Algébriquement cette expression est la moyenne des valeurs des génotypes contenant l'association  $A_i A_j A_k$ . Pour obtenir l'effet spécifique à l'association  $A_i A_j A_k$  il faut retrancher la moyenne générale, les effets principaux et les interactions du premier ordre.

$$\begin{aligned}
 A_i A_j A_k \left( \sum_s p_s A_s \right) &= \alpha_i + \alpha_j + \alpha_k + \beta_{ij} + \beta_{ik} + \beta_{jk} + \mu = \\
 A_i A_j A_k \left( \sum_s p_s A_s \right) &= (A_i - \sum_s p_s A_s) \left( \sum_s p_s A_s \right)^3 + (A_j - \sum_s p_s A_s) \left( \sum_s p_s A_s \right)^3 \\
 &\quad - (A_k - \sum_s p_s A_s) \left( \sum_s p_s A_s \right)^3 - \left( \sum_s p_s A_s \right)^2 (A_i - \sum_s p_s A_s) (A_j - \sum_s p_s A_s) \\
 &\quad - \left( \sum_s p_s A_s \right)^2 (A_i - \sum_s p_s A_s) (A_k - \sum_s p_s A_s) \\
 &\quad - \left( \sum_s p_s A_s \right)^2 (A_j - \sum_s p_s A_s) (A_k - \sum_s p_s A_s) - \left( \sum_s p_s A_s \right)^4 = \\
 \left( \sum_s p_s A_s \right) (A_i - \sum_s p_s A_s) (A_j - \sum_s p_s A_s) (A_k - \sum_s p_s A_s) &= \rho_{ijk}^{\text{UVU}}
 \end{aligned}$$

on voit que  $\gamma_{ijk}$  représente un écart propre à l'association  $A_i A_j A_k$

$\gamma_{ijk}$  est l'effet trigénique associé aux gènes  $A_i A_j A_k$

c'est l'équivalent d'une interaction de second ordre.

$$4^\circ) \text{ Posons } \delta_{ijkl} = \rho_{ijkl}^{\text{UVU}}$$

$\delta_{ijkl}$  représente un écart propre à l'association  $A_i A_j A_k A_l$

$\delta_{ijkl}$  est l'effet tétragénique associé aux gènes  $A_i A_j A_k A_l$

c'est l'équivalent d'une interaction de troisième ordre.

#### b) - Expression de la variance génétique

$y_{ijkl}$  mesure du génotype  $A_i A_j A_k A_l$  se décompose en l'expression suivante :

$$\begin{aligned}
 y_{ijkl} &= \mu + \alpha_i + \alpha_j + \alpha_k + \alpha_l + \beta_{ij} + \beta_{ik} + \beta_{jk} + \beta_{jl} + \beta_{kl} + \gamma_{ijk} \\
 &\quad + \gamma_{ijl} + \gamma_{ikl} + \gamma_{jkl} + \delta_{ijkl}
 \end{aligned}$$

la variance génétique est égale à la variance de  $y_{ijkl}$

on peut supposer que l'ordre des gènes d'un individu quelconque est aléatoire :

La population est en panmixie à l'équilibre.

$$\text{Prob}\{\text{gène } A_i \text{ soit un gène donné } A_r\} = p_r$$

c'est-à-dire la probabilité que le premier gène au locus considéré soit un gène  $A_r$  particulier est  $p_r$  fréquence du gène  $A_r$  dans la population.

Les gènes possédés par un individu pris au hasard dans la population sont aléatoires et nous pouvons valablement parler d'espérance mathématique :

$$\begin{aligned} E(\alpha_i) &= \sum_i p_i \alpha_i \\ &= \sum_i p_i (A_i - \rho) \rho^3 \\ &= \rho^3 \left( \sum_i p_i A_i - \rho \sum_i p_i \right) \end{aligned}$$

$$\text{mais } \sum_i p_i = 1 \quad \text{et} \quad \sum_i p_i A_i = \rho$$

$$\text{donc } E(\alpha_i) = 0$$

de la même façon on montre que

$$E(\beta_{ij}) = \sum_{ij} p_i p_j \beta_{ij} = 0$$

$$E(\gamma_{ijk}) = \sum_{ijk} p_i p_j p_k \gamma_{ijk} = 0$$

$$E(\delta_{ijkl}) = \sum_{ijkl} p_i p_j p_k p_l \delta_{ijkl} = 0$$

de même

$$E(\alpha_i \alpha_j) = E(\alpha_i \beta_{ij}) = E(\beta_{ij} \beta_{jk}) = \dots = 0$$

par exemple

$$\begin{aligned} E(\alpha_i \beta_{ij}) &= \sum_{ij} p_i p_j \alpha_i \beta_{ij} \\ &= \sum_{ij} p_i p_j (A_i - \rho) \rho^3 (A_j - \rho) \rho^2 \\ &= \rho^5 \sum_i p_i (A_i - \rho)^2 \left[ \sum_j p_j (A_j - \rho) \right] = 0 \end{aligned}$$

par conséquent la variance génétique s'écrit :

$$\begin{aligned}\sigma_G^2 &= \text{var}(y_{ijkl}) = E\{(y_{ijkl} - \mu)^2\} \\ &= E\{(\alpha_i + \alpha_j + \alpha_k + \alpha_l + \beta_{ij} + \beta_{ik} + \beta_{il} + \beta_{jk} + \beta_{jl} + \beta_{kl} \\ &\quad + \gamma_{ijk} + \gamma_{ijl} + \gamma_{ikl} + \gamma_{jkl} + \delta_{ijkl})^2\} \\ &= E(\alpha_i^2) + E(\alpha_j^2) + E(\alpha_k^2) + E(\alpha_l^2) + E(\beta_{ij}^2) + E(\beta_{ik}^2) + E(\beta_{il}^2) \\ &\quad + E(\beta_{jk}^2) + E(\beta_{jl}^2) + E(\beta_{kl}^2) + E(\gamma_{ijk}^2) + E(\gamma_{ijl}^2) + E(\gamma_{ikl}^2) \\ &\quad + E(\gamma_{jkl}^2) + E(\delta_{ijkl}^2)\end{aligned}$$

$\sigma_G^2$  est une somme d'espérances de carrés des effets car toutes les espérances mathématiques des termes produits sont nulles.

On peut poser :

$$E(\alpha_i^2) = \frac{1}{4} \sigma_A^2 ; \sigma_A^2 \text{ variance additive}$$

$$E(\beta_{ij}^2) = \frac{1}{6} \sigma_D^2 ; \sigma_D^2 \text{ variance digénique}$$

$$E(\gamma_{ijk}^2) = \frac{1}{4} \sigma_T^2 ; \sigma_T^2 \text{ variance trigénique}$$

$$E(\delta_{ijkl}^2) = \sigma_Q^2 ; \sigma_Q^2 \text{ variance tétragénique}$$

On obtient :

$$\sigma_G^2 = \sigma_A^2 + \sigma_D^2 + \sigma_T^2 + \sigma_Q^2$$

## 2°) Cas de n loci

### a) définition des effets géniques

Le génotype d'un individu tétraploïde à n loci s'écrit :

$$A_i^1 A_j^1 A_k^1 A_l^1 A_i^2 A_j^2 A_k^2 A_l^2 \dots \dots \dots A_i^n A_j^n A_k^n A_l^n$$

où  $A_i^1$  désigne un allèle du premier locus

on peut écrire ce génotype sous une forme de produit symbolique

$$\prod_{a=1}^n A_i^a A_j^a A_k^a A_l^a \quad \text{"a" désignant le locus}$$

En désignant par  $p_i^a$  la fréquence de l'allèle  $A_i^a$  dans la population la composition génotypique de la population peut s'écrire symboliquement

$$\prod_{a=1}^n \left( \sum_i p_i^a A_i^a \right)^4 = \sum_{a=1}^n \left[ \prod_{i,j,k,l} p_i^a p_j^a p_k^a p_l^a A_i^a A_j^a A_k^a A_l^a \right]$$

comme précédemment

en remplaçant chaque génotype par sa contribution dans la mesure du caractère l'expression devient algébrique et représente la moyenne  $\mu$  de la population panmictique en posant

$$\rho^a = \sum_i p_i^a A_i^a$$

$$\mu = \prod_{a=1}^n (\rho^a)^4$$

soit

$$v_i^a = A_i^a - \rho^a$$

$v_i^a$  représente un écart de l'effet du gène  $A_i^a$  à la moyenne  $\rho^a$  des effets des allèles du locus "a".

$A_i^a$  est donc composé d'un effet moyen commun à tous les allèles du locus "a" et d'un effet propre au gène  $A_i^a$

on peut écrire

$$\prod_{a=1}^n A_i^a A_j^a A_k^a A_l^a = \prod_{a=1}^n (\rho^a + v_i^a)(\rho^a + v_j^a)(\rho^a + v_k^a)(\rho^a + v_l^a)$$

et comme pour un locus on peut définir les différents termes du développement de ce produit de la façon suivante :

$$\mu = \prod_{a=1}^n (\rho^a)^4$$

moyenne génotypique de la population

$$\alpha_i^a = (\rho^a)^3 \nu_i^a \prod_{a \neq a'} (\rho^{a'})^4$$

effet additif du gène  $\Lambda_i^a$  ou effet principal

$$\beta_{ij}^a = (\rho^a)^2 \nu_i^a \nu_j^a \prod_{a \neq a'} (\rho^{a'})^4$$

effet digénique associé aux gènes  $\Lambda_i^a$  et  $\Lambda_j^a$  ou interaction du premier ordre

$$\gamma_{ijk}^a = \rho^a \nu_i^a \nu_j^a \nu_k^a \prod_{a \neq a'} (\rho^{a'})^4$$

effet trigénique associé aux gènes  $\Lambda_i^a$   $\Lambda_j^a$   $\Lambda_k^a$  ou interaction du deuxième ordre

$$\delta_{ijkl}^a = \nu_i^a \nu_j^a \nu_k^a \nu_l^a \prod_{a \neq a'} (\rho^{a'})^4$$

effet quadrigénique associé aux gènes  $\Lambda_i^a$   $\Lambda_j^a$   $\Lambda_k^a$   $\Lambda_l^a$  ou interaction du troisième ordre

$$(\alpha^a \alpha^{a'})_{i,j} = (\rho^a)^3 (\rho^{a'})^3 \nu_i^a \nu_j^{a'} \prod_{a'' \neq a, a'} (\rho^{a''})^4$$

effet d'interaction additif par additif

$$(\beta^a \alpha^{a'})_{ij,k} = (\rho^a)^2 (\rho^{a'})^3 \nu_i^a \nu_j^a \nu_k^{a'} \prod_{a'' \neq a, a'} (\rho^{a''})^4$$

effet d'interaction digénique par additif

$$(\beta^a \beta^{a'})_{ij,kl} \quad \text{effet d'interaction digénique par digénique}$$

etc.....

$$\prod_{a=1}^n \Lambda_i^a \Lambda_j^a \Lambda_k^a \Lambda_l^a = \mu + \sum_a (\alpha_i^a + \alpha_j^a + \alpha_k^a + \alpha_l^a) + \sum_a (\beta_{ij}^a + \beta_{ik}^a + \dots)$$

$$+ \sum_a (\gamma_{ijk}^a + \gamma_{ijl}^a + \dots) + \sum_a \delta_{ijkl}^a + \sum_{a, a'} [(\alpha^a \alpha^{a'})_{i,j} + \dots]$$

$$+ \sum_{a, a'} [(\beta^a \alpha^{a'})_{ij,k} + \dots] + \sum_{a, a'} [(\beta^a \beta^{a'})_{ij,kl} + \dots]$$

$$+ \dots$$

Comme précédemment on montre que :

$$E(\alpha_i^a) = 0$$

$$E(\alpha_i^a \alpha_j^a) = 0$$

$$E(\alpha_i^a \alpha_j^{a'}) = 0$$

etc .....

b) - Expression de la variance génétique

En posant :

$$E\{(\alpha_i^a)^2\} = \frac{1}{4} \sigma_A^2(a)$$

$$E\{(\beta_{ij}^a)^2\} = \frac{1}{6} \sigma_D^2(a)$$

$$E\{(\gamma_{ijk}^a)^2\} = \frac{1}{4} \sigma_T^2(a)$$

$$E\{(\delta_{ijkl}^a)^2\} = \sigma_Q^2(a)$$

$$E\{(\alpha_i^a \alpha_j^{a'})_{i,j}^2\} = \frac{1}{16} \sigma_{AA}^2(aa') \quad \text{etc .....$$

et en sommant sur tous les loci l'espérance mathématique des produits étant nulle :

$$\sigma_A^2 = \sum_a \sigma_A^2(a) \quad \text{variance additive totale}$$

$$\sigma_D^2 = \sum_a \sigma_D^2(a) \quad \text{variance digénique totale}$$

$$\sigma_T^2 = \sum_a \sigma_T^2(a) \quad \text{variance trigénique totale}$$

$$\sigma_Q^2 = \sum_a \sigma_Q^2(a) \quad \text{variance tétragénique totale}$$

$$\sigma_{AA}^2 = \sum_{aa'} \sigma_{AA}^2(aa') \quad \text{variance additive x additive totale}$$

etc ...

La variance de la valeur génotypique se décomposera comme suit :

$$\sigma_G^2 = \sigma_A^2 + \sigma_D^2 + \sigma_T^2 + \sigma_Q^2 + \sigma_{AA}^2 + \sigma_{AD}^2 + \sigma_{AT}^2 + \sigma_{AQ}^2 + \sigma_{DD}^2 + \sigma_{DT}^2 + \sigma_{DQ}^2 + \\ \sigma_{TT}^2 + \sigma_{TQ}^2 + \sigma_{QQ}^2 + \sigma_{AAA}^2 + \sigma_{AAD}^2 + \sigma_{AAT}^2 + \sigma_{AAQ}^2 + \sigma_{ADQ}^2 + \dots\dots\dots$$

D - Décomposition des covariances entre individus apparentés en fonction des composants de la variance génétique

Le problème qui reste à traiter est celui de la décomposition des covariances entre individus apparentés en fonction des différents composants de  $\sigma_G^2$  qui viennent d'être définis.

La méthode a été donnée par G.MALECOT (1948) et par O.KEMPTHORPE (1957).

1°) Cas d'un locus

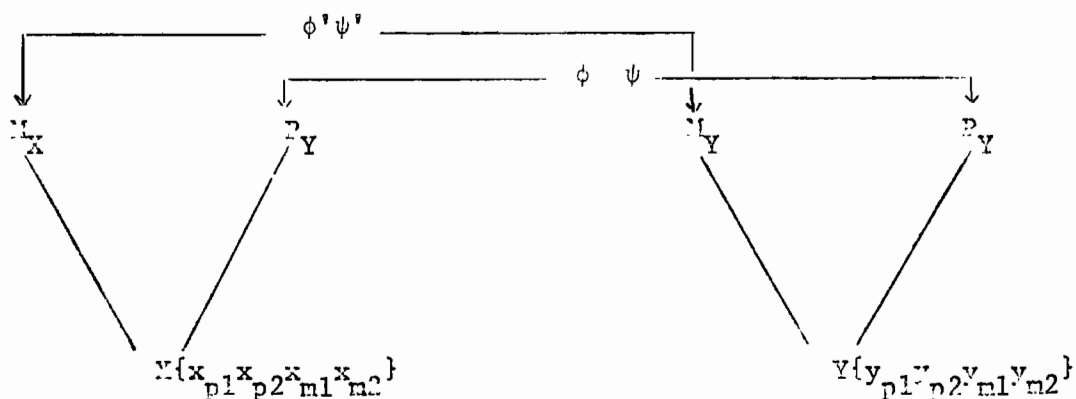
soient deux individus tétraploïdes X et Y

non consanguins liés par une relation de parenté telle que la moitié des gènes viennent d'une source mâle et l'autre d'une source femelle indépendante de la première.

Le génotype de X est  $x_{p1}x_{p2} \ x_{m1}x_{m2}$

Le génotype de Y est  $y_{p1}y_{p2} \ y_{m1}y_{m2}$

Les  $x_p$  et  $y_p$  proviennent de la source mâle (source paternelle) les  $x_m$  et  $y_m$  proviennent de la source femelle (source maternelle).





soit  $A_{p1} A_{p2} \dots A_{pn}$  la série allélique transmise par la voie paternelle

et  $Q_{pi} = \text{Prob.}\{X \text{ et } Y \text{ reçoivent par voie paternelle un gène donné } A_{pi}\}$

on pose  $\phi = \sum Q_{pi}$

$\sum$  est la somme étendue à tous les allèles de la série mâle

$\phi$  représente le coefficient de simple parenté (mâle) c'est-à-dire la probabilité de tirer au hasard un gène mâle de X,  $A_i^X$ , et un gène mâle de Y,  $A_j^Y$  tels que  $A_i^X$  soit identique (au sens de MALECOT) à  $A_j^Y$  soit  $A_{m1}, A_{m2} \dots A_{mr}$  la série allélique transmise par la voie maternelle

et  $Q_{mi} = \text{prob.}\{X \text{ et } Y \text{ reçoivent par voie maternelle un gène donné } A_{mi}\}$

on pose  $\phi' = \sum Q_{mi}$ .

$\sum$  est la somme étendue à tous les allèles de la série femelle

$\phi'$  représente le coefficient de simple parenté (femelle) c'est-à-dire la probabilité de tirer au hasard un gène femelle de X,  $A_i^X$  et un gène femelle Y,  $A_j^Y$ , tels que  $A_i^X$  soit identique à  $A_j^Y$

soit  $R_{pipj} = \text{Prob.}\{X \text{ et } Y \text{ reçoivent par voie paternelle un couple de gènes donné } A_{pi} A_{pj}\}$

et  $\psi = \sum R_{pipj}$

soit  $R'_{mimj} = \text{Prob.}\{X \text{ et } Y \text{ reçoivent par voie maternelle un couple de gènes donné } A_{mi} A_{mj}\}$

et  $\psi' = \sum R'_{mimj}$

$\psi$  et  $\psi'$  représentent respectivement les coefficients de double parenté (mâle et femelle) c'est-à-dire la probabilité de tirer un couple de gènes mâles (ou femelle) de X,  $A_i^X A_j^X$  et un couple de gènes mâles (ou femelle) de Y,  $A_k^Y A_l^Y$  tels que  $A_i^X A_j^X$  soit identique à  $A_k^Y A_l^Y$ .

$\phi, \phi', \psi, \psi'$  sont des coefficients qui déterminent le lien de parenté entre les parents de X et Y et par conséquent le lien de parenté entre X et Y eux mêmes.

La décomposition de la covariance entre les valeurs génotypiques de X et Y va être faite en fonction de ces coefficients.

Soient  $G_X$  et  $G_Y$  les valeurs génotypiques de X et Y. On désigne généralement par  $\text{cov}(XY)$  la covariance entre  $G_X$  et  $G_Y$ .

$$\begin{aligned} G_X = & \mu_X + \alpha_{xp_1} + \alpha_{xp_2} + \alpha_{xm_1} + \alpha_{xm_2} + \beta_{xp_1xp_2} + \beta_{xp_1xm_1} + \beta_{xp_1xm_2} \\ & + \beta_{xp_2xm_1} + \beta_{xp_2xm_2} + \beta_{xm_1xm_2} + \gamma_{xp_1xp_2xm_1} + \gamma_{xp_1xp_2xm_2} \\ & + \gamma_{xp_1xm_1xm_2} + \gamma_{xp_2xm_1xm_2} + \delta_{xp_1xp_2xm_1xm_2} \end{aligned}$$

$\alpha_{xp_1}$  représente l'effet additif du gène  $xp_1$  ; un des deux gènes reçus par X de son père.

$\beta_{xp_1xp_2}$  représente l'effet digénique spécifique aux gènes  $xp_1$  et  $xp_2$  (ou effet d'interaction entre les gènes  $xp_1$  et  $xp_2$ ).

etc .....

Pour simplifier on peut écrire  $G_X$  sous la forme :

$$G_X = \mu_X + \sum \alpha_{x_i} + \sum \beta_{x_i x_j} + \sum \gamma_{x_i x_j x_k} + \delta_{x_i x_j x_k x_l}$$

et de la même façon pour  $G_Y$

$$G_Y = \mu_Y + \sum \alpha_{y_i} + \sum \beta_{y_i y_j} + \sum \gamma_{y_i y_j y_k} + \delta_{y_i y_j y_k y_l}$$

$$\begin{aligned} \text{cov}(XY) &= E\{(G_X - \mu_X)(G_Y - \mu_Y)\} \\ &= \sum E\{\alpha_{x_i} \alpha_{y_j}\} + \sum E\{\alpha_{x_i} \beta_{y_j y_k}\} + \sum E\{\alpha_{x_i} \gamma_{y_j y_k y_l}\} + \sum E\{\alpha_{x_i} \delta_{y_j y_k y_l y_m}\} \\ &+ \sum E\{\beta_{x_i x_j} \alpha_{y_k}\} + \sum E\{\beta_{x_i x_j} \beta_{y_l y_m}\} + \sum E\{\beta_{x_i x_j} \gamma_{y_k y_l y_m}\} \end{aligned}$$

$$\begin{aligned}
 & + \sum E\{\beta_{x_i x_j} \delta_{y_k y_l y_m y_n}\} + \sum E\{\gamma_{x_i x_j x_k} \alpha_{y_l}\} + \sum E\{\gamma_{x_i x_j x_k} \beta_{y_l y_m}\} \\
 & + \sum E\{\gamma_{x_i x_j x_k} \gamma_{y_l y_m y_n}\} + \sum E\{\gamma_{x_i x_j x_k} \delta_{y_l y_m y_n y_p}\} \\
 & + \sum E\{\delta_{x_i x_j x_k x_l} \alpha_{y_m}\} + \sum E\{\delta_{x_i x_j x_k x_l} \beta_{y_m y_n}\} \\
 & + \sum E\{\delta_{x_i x_j x_k x_l} \gamma_{y_m y_n y_p}\} + \sum E\{\delta_{x_i x_j x_k x_l} \delta_{y_m y_n y_p y_q}\}
 \end{aligned}$$

Mais nous avons vu au paragraphe 3 que pour  $i \neq j$  :

$$E\{\alpha_i \alpha_j\} = 0 \quad E\{\alpha_i \beta_{ij}\} = 0 \text{ et pour } i \neq j, i \neq k, j \neq k : E\{\beta_{ij} \beta_{jk}\} = 0 \text{ etc...}$$

$$\begin{aligned}
 \sum E\{\alpha_{x_i} \alpha_{y_j}\} &= \sum \text{Prob.}(x_i = y_j) E\{\alpha_{x_i}^2\} \\
 &= \sum [\text{Prob}(x_p = y_p) + \text{Prob}(x_m = y_m)] \frac{1}{4} \sigma_A^2
 \end{aligned}$$

puisque les gènes de source paternelle sont indépendants des gènes de source maternelle.

$$\sum E\{\alpha_{x_i} \alpha_{y_j}\} = (\phi + \phi') \frac{1}{4} \sigma_A^2$$

d'après la remarque ci-dessus  $\sum E\{\alpha_{x_i} \beta_{y_j y_k}\} = 0$

$$\sum E\{\alpha_{x_i} \gamma_{y_j y_k y_l}\} = 0$$

$$\sum E\{\alpha_{x_i} \delta_{y_j y_k y_l y_m}\} = 0$$

$$\sum E\{\beta_{x_i x_j} \beta_{y_k y_l}\} = \sum \text{Prob}(x_i x_j = y_k y_l) E\{\beta_{x_i x_j}^2\}$$

$$\begin{aligned}
 &= [\sum \text{Prob}(x_p x_m = y_p y_m) + \text{Prob}(x_{p1} x_{p2} = y_{p1} y_{p2}) \\
 &\quad + \text{Prob}(x_{m1} x_{m2} = y_{m1} y_{m2})] \frac{1}{6} \sigma_D^2
 \end{aligned}$$

$$\sum E\{\beta_{x_i x_j} \beta_{y_k y_l}\} = (\phi \phi' + \psi + \psi') \frac{1}{6} \sigma_D^2$$

$$\sum E\{\beta_{x_i x_j} \gamma_{y_k y_l y_m}\} = 0$$

$$\sum E\{\beta_{x_i x_j} \delta_{y_k y_l y_m y_n}\} = 0$$

$$\sum E\{\gamma_{x_i x_j x_k} \gamma_{y_l y_m y_n}\} = [\sum \text{Prob}(x_p = y_p, x_{m_1} x_{m_2} = y_{m_1} y_{m_2}) +$$

$$\sum \text{Prob}(y_m = y_m, x_{p_1} x_{p_2} = y_{p_1} y_{p_2})] E\{\gamma_{x_i x_j x_k}^2\}$$

$$= [\sum \text{Prob}(x_p = y_p) \text{Prob}(x_{m_1} x_{m_2} = y_{m_1} y_{m_2}) +$$

$$\sum \text{Prob}(x_m = y_m) \text{Prob}(x_{p_1} x_{p_2} = y_{p_1} y_{p_2})] \frac{1}{4} \sigma_T^2$$

$$\sum E\{\gamma_{x_i x_j x_k} \gamma_{y_l y_m y_n}\} = (\phi \psi' + \phi' \psi) \frac{1}{4} \sigma_T^2$$

$$\sum E\{\gamma_{x_i x_j x_k} \delta_{y_l y_m y_n y_p}\} = 0$$

$$E\{\delta_{x_i x_j x_k x_l} \delta_{y_m y_n y_p y_q}\} = E\{\delta_{x_{p_1} x_{p_2} x_{m_1} x_{m_2}} \delta_{y_{p_1} y_{p_2} y_{m_1} y_{m_2}}\}$$

$$= [\text{Prob}(x_{p_1} x_{p_2} = y_{p_1} y_{p_2}) \text{Prob}(x_{m_1} x_{m_2} = y_{m_1} y_{m_2})] \sigma_Q^2$$

$$E\{\delta_{x_i x_j x_k x_l} \delta_{y_m y_n y_p y_q}\} = \psi \psi' \sigma_Q^2$$

d'où finalement

$$\text{cov}(XY) = (\phi + \phi') \frac{\sigma_A^2}{4} + (\phi \phi' + \psi + \psi') \frac{\sigma_D^2}{6} + (\phi \psi' + \phi' \psi) \frac{\sigma_T^2}{4} + \psi \psi' \sigma_Q^2$$

## 2°) Cas de deux loci

Si on considère un caractère dépendant de deux loci les génotypes de deux individus X et Y, pour lesquels on fait les mêmes hypothèses que pour le cas d'un locus s'écrivent respectivement :

$$x_{i1}x_{j1}x_{k1}x_{11} \quad x_{i2}x_{j2}x_{k2}x_{12}$$

$$y_{i1}y_{j1}y_{k1}y_{11} \quad y_{i2}y_{j2}y_{k2}y_{12}$$

$$\begin{aligned} G_X = & \mu_X + \alpha_{x_{i1}} + \alpha_{x_{j1}} + \alpha_{x_{k1}} + \alpha_{x_{11}} + \alpha_{x_{i2}} + \alpha_{x_{j2}} + \alpha_{x_{k2}} + \alpha_{x_{12}} + \beta_{x_{i1}x_{j1}} \\ & + \beta_{x_{i1}x_{k1}} + \beta_{x_{i1}x_{11}} + \beta_{x_{j1}x_{k1}} + \beta_{x_{j1}x_{11}} + \beta_{x_{k1}x_{11}} + \beta_{x_{i2}x_{j2}} + \beta_{x_{i2}x_{k2}} \\ & + \beta_{x_{i2}x_{12}} + \beta_{x_{j2}x_{k2}} + \beta_{x_{j2}x_{12}} + \beta_{x_{k2}x_{12}} + \gamma_{x_{i1}x_{j1}x_{k1}} + \gamma_{x_{i1}x_{k1}x_{11}} \\ & + \gamma_{x_{i1}x_{j1}x_{11}} + \gamma_{x_{j1}x_{k1}x_{11}} + \gamma_{x_{i2}x_{j2}x_{k2}} + \gamma_{x_{i2}x_{j2}x_{12}} + \gamma_{x_{i2}x_{k2}x_{12}} \\ & + \gamma_{x_{j2}x_{k2}x_{12}} + \delta_{x_{i1}x_{j1}x_{k1}x_{11}} + \delta_{x_{i2}x_{j2}x_{k2}x_{12}} \\ & + (\alpha\alpha)_{x_{i1}x_{i2}} + (\alpha\alpha)_{x_{i1}x_{j2}} + \dots \\ & + (\alpha\beta)_{x_{i1}x_{i2}x_{j2}} + (\alpha\beta)_{x_{i1}x_{i2}x_{k2}} + \dots \\ & + (\alpha\gamma)_{x_{i1}x_{i2}x_{j2}x_{k2}} + (\alpha\gamma)_{x_{i1}x_{i2}x_{j2}x_{12}} + \dots \\ & + (\alpha\delta)_{x_{i1}x_{i2}x_{j2}x_{k2}x_{12}} + (\alpha\delta)_{x_{j1}x_{i2}x_{j2}x_{k2}x_{12}} + \dots \\ & + (\beta\beta)_{x_{i1}x_{j1}x_{i2}x_{j2}} + (\beta\beta)_{x_{i1}x_{j1}x_{i2}x_{k2}} + \dots \\ & + (\beta\gamma)_{x_{i1}x_{j1}x_{i2}x_{j2}x_{k2}} + (\beta\gamma)_{x_{i1}x_{j1}x_{i2}x_{j2}x_{12}} + \dots \\ & + (\beta\delta)_{x_{i1}x_{j1}x_{i2}x_{j2}x_{k2}x_{12}} + \dots + (\gamma\gamma)_{x_{i1}x_{j1}x_{k1}x_{i2}x_{j2}x_{k2}} + \dots \\ & + (\gamma\delta)_{x_{i1}x_{j1}x_{k1}x_{i2}x_{j2}x_{k2}x_{12}} + \dots + (\delta\delta)_{x_{i1}x_{j1}x_{k1}x_{11}x_{i2}x_{j2}x_{k2}x_{12}} \end{aligned}$$

la formule est analogue pour  $G_Y$

$$\text{cov}(XY) = E\{(G_X - \mu_X)(G_Y - \mu_Y)\}$$

on va examiner successivement les différents termes de  $\text{Cov}(XY)$  en sommant les termes analogues

$$\sum E\{\alpha_{x_{ia}} \alpha_{y_{ia}}\} = \sum \text{Prob}(x_{ia} = y_{ia}) E\{(\alpha^i)^2\}$$

cette expression a été calculée dans le cas d'un locus on a donc

$$\begin{aligned} \sum E\{\alpha_{x_{i1}} \alpha_{y_{j1}}\} + \sum E\{\alpha_{x_{i2}} \alpha_{y_{j2}}\} &= (\phi + \phi') [E\{(\alpha^1)^2\} + E\{(\alpha^2)^2\}] \\ &= (\phi + \phi') \left[ \frac{1}{4} \sigma_A^2(1) + \frac{1}{4} \sigma_A^2(2) \right] \\ &= (\phi + \phi') \frac{1}{4} \sigma_A^2 \end{aligned}$$

Tous les termes comprenant des paramètres des deux loci différents sont nuls puisque les deux loci ont des séries alléliques indépendantes. Pour un même locus on a vu précédemment que les espérances mathématiques des termes contenant des paramètres différents (tels que  $\alpha\beta, \alpha\gamma, \alpha\delta$  etc.....) sont nulles.

Les termes suivants non nuls sont ceux des sommes :

$$\sum E\{\beta_{x_{ia}x_{ja}} \beta_{y_{ia}y_{ja}}\} \quad (\text{pour } a = 1 \quad a = 2)$$

qui ont été calculées dans le cas d'un locus.

on a donc :

$$\begin{aligned} \sum E\{\beta_{x_{i1}x_{j1}} \beta_{y_{i1}y_{j1}}\} + \sum E\{\beta_{x_{i2}x_{j2}} \beta_{y_{i2}y_{j2}}\} \\ &= (\phi\phi' + \psi + \psi') [E\{(\beta^1)^2\} + E\{(\beta^2)^2\}] \\ &= (\phi\phi' + \psi + \psi') \left[ \frac{1}{6} \sigma_D^2(1) + \frac{1}{6} \sigma_D^2(2) \right] \\ &= (\phi\phi' + \psi + \psi') \frac{1}{6} \sigma_D^2 \end{aligned}$$

Pour les autres termes de la dominance le calcul est analogue ; ils se décomposent en une variance pour le premier locus et une variance pour le second que l'on réunit en une variance totale comme on l'a défini dans le paragraphe 2.2°.

Pour les termes de l'épistasie on a comme précédemment des termes nuls. Les termes comprenant des paramètres des deux loci différents sont nuls, les termes comprenant des paramètres différents sont nuls.

$$\begin{aligned} \sum E\{(\alpha\alpha)_{x_{i1}x_{i2}} (\alpha\alpha)_{y_{i1}y_{i2}}\} &= \sum \text{Prob}(x_{i1} = y_{i1}) \times \sum \text{Prob}(x_{i2} = y_{i2}) \times E\{(\alpha^1\alpha^2)\} \\ &= (\phi + \phi')(\phi + \phi') \frac{1}{16} \sigma_{\Lambda\Lambda}^2 \end{aligned}$$

$$\sum E\{(\alpha\beta)_{x_{i1}x_{i2}x_{j2}} (\alpha\beta)_{y_{i1}y_{i2}y_{j2}}\} + \sum E\{(\alpha\beta)_{x_{i2}x_{i1}x_{j1}} (\alpha\beta)_{y_{i2}y_{i1}y_{j1}}\} =$$

$$\sum \text{Prob}(x_{i1} = y_{i1}) \sum \text{Prob}(x_{i2}x_{j2} = y_{i2}y_{j2}) E\{(\alpha^1\beta^2)^2\} +$$

$$\sum \text{Prob}(x_{i2} = y_{i2}) \sum \text{Prob}(x_{i1}x_{j1} = y_{i1}y_{j1}) E\{(\alpha^2\beta^1)^2\} =$$

$$(\phi + \phi')(\phi\phi' + \psi + \psi') [E\{(\alpha^1\beta^2)^2\} + E\{(\alpha^1\beta^2)^2\}] =$$

$$(\phi + \phi')(\phi\phi' + \psi + \psi') \left[ \frac{1}{24} \sigma_{A_1D_2}^2 + \frac{1}{24} \sigma_{A_2D_1}^2 \right] =$$

$$(\phi + \phi')(\phi\phi' + \psi + \psi') \frac{1}{24} \sigma_{\Lambda D}^2$$

et en faisant de même pour toutes les sommes on obtient :

$$\begin{aligned} \text{cov}(XY) &= (\phi + \phi') \frac{\sigma_A^2}{4} + (\phi\phi' + \psi + \psi') \frac{\sigma_D^2}{6} + (\phi\psi' + \phi'\psi) \frac{\sigma_T^2}{4} + \psi\psi' \sigma_Q^2 \\ &+ (\phi + \phi')^2 \frac{\sigma_{\Lambda\Lambda}^2}{16} + (\phi\phi' + \psi + \psi')(\phi + \phi') \frac{\sigma_{\Lambda D}^2}{24} + (\phi\psi' + \phi'\psi)(\phi + \phi') \frac{\sigma_{\Lambda T}^2}{16} \\ &+ (\phi + \phi')\psi\psi' \frac{\sigma_{\Lambda Q}^2}{4} + (\phi\phi' + \psi + \psi')^2 \frac{\sigma_{DD}^2}{36} + (\phi\phi' + \psi + \psi')(\phi\psi' + \phi'\psi) \frac{\sigma_{DT}^2}{24} \end{aligned}$$

$$+ (\phi\phi' + \psi + \psi')\psi\psi' \frac{\sigma_{DQ}^2}{6} + (\phi\psi' + \phi'\psi)^2 \frac{\sigma_{TT}^2}{16} + \psi\psi' (\phi\psi' + \phi'\psi) \frac{\sigma_{TQ}^2}{4} + \psi\psi'^2 \sigma_{QQ}^2$$

La formule se généralise sous la forme suivante pour  $n$  loci :

$$\text{cov}(XY) = \sum_{\substack{0 < r, s, p, q < n \\ 1 < r+s+p+q < n}} \left(\frac{\phi+\phi'}{4}\right)^r \left(\frac{\phi\phi' + \psi + \psi'}{6}\right)^s \left(\frac{\phi\psi' + \phi'\psi}{4}\right)^p (\psi\psi')^q \sigma_{AD}^2 r_D s_T p_Q q$$

### E - Calcul des coefficients ( $\phi, \phi', \psi$ et $\psi'$ ) de parenté

Soit  $F$  le coefficient de consanguinité de la population étudiée. Les coefficients de parenté  $\phi, \psi, \phi'$  et  $\psi'$  dépendent évidemment de  $F$ . On étudiera le cas où  $F = 0$ .

On étudiera successivement les différentes relations de parenté existant dans la table diallèle.

#### Relation parent-enfant

Les génotypes de  $X$  et  $Y$ ,  $X$  étant le père de  $Y$ , sont respectivement :

$$(x_p x_p' \ x_m x_m') \text{ et } (y_p y_p' \ y_m y_m')$$

$$\phi = \text{Prob}(x_p = y_p) + \text{Prob}(x_p' = y_p) + \text{Prob}(x_p = y_p') + \text{Prob}(x_p' = y_p')$$

$$+ \text{Prob}(x_m = y_p) + \text{Prob}(x_m' = y_p) + \text{Prob}(x_m = y_p') + \text{Prob}(x_m' = y_p')$$

$$\text{Prob}(x_p = y_p) = \frac{1}{4} \text{ puisque } y_p \text{ est l'un des quatre gènes } x_p, x_p', x_m, x_m'.$$

Il en est de même pour toutes les probabilités qui composent  $\phi$  donc

$$\phi = 3 \times \frac{1}{4} \quad \phi = \frac{3}{4}$$

On a supposé que la mère de  $X$  et la mère de  $Y$  sont indépendantes donc  $\phi' = 0$

$$\psi = \text{Prob}(x_p x_p' = y_p y_p') + \text{Prob}(x_p x_m = y_p y_p') + \text{Prob}(x_p x_m' = y_p y_p') +$$

$$\text{Prob}(x_p' x_m = y_p y_p') + \text{Prob}(x_p' x_m' = y_p y_p') + \text{Prob}(x_m x_m' = y_p y_p')$$

$\text{Prob}(x_p x_p' = y_p y_p') = \frac{1}{6}$  puisque  $x_p x_p'$  est l'un des six couples de gènes pris parmi  $x_p, x_p', x_m, x_m'$ .



$$\psi = 1.$$

$\phi'$  étant nul  $\psi'$  l'est aussi.

d'où finalement  $\phi = 2$ ,  $\psi = 1$ ,  $\phi' = 0$ ,  $\psi' = 0$ .

#### Relation entre frères

X et Y sont frères

Soit abcd le génotype du père de X et Y

$$\begin{aligned} \text{Prob}(x_p = y_p) &= \text{Prob}(x_p = a, y_p = a) + \text{Prob}(x_p = b, y_p = b) + \text{Prob}(x_p = c, y_p = c) \\ &\quad + \text{Prob}(x_p = d, y_p = d) \\ &= \left(\frac{1}{4} \times \frac{1}{4}\right) \times 4 = \frac{1}{4} \end{aligned}$$

Il en est de même pour les autres probabilités donc  $\phi = 1$ .

La relation de parenté étant symétrique :  $\phi' = 1$ .

$$\psi = \text{Prob}(x_p x'_p = y_p y'_p)$$

Il existe 36 couples  $x_p x'_p$ ,  $y_p y'_p$  possibles qui peuvent être (ab, ab) (ab, ac)

(ad, bc) etc ..... dont 6 sont tels que  $x_p x'_p = y_p y'_p$  donc :

$$\psi = \frac{6}{36} = \frac{1}{6} \text{ et bien sûr } \psi' = \frac{1}{6}.$$

d'où finalement  $\phi = 1$ ,  $\psi = \frac{1}{6}$ ,  $\phi' = 1$ ,  $\psi' = \frac{1}{6}$ .

#### Relation demi-frères

X et Y sont ici demi-frères

Le calcul est analogue au cas précédent pour  $\phi$  et  $\psi$  ;  $\phi = 1$ ,  $\psi = \frac{1}{6}$

Les mères sont indépendantes  $\phi' = \psi' = 0$

d'où le tableau récapitulatif des coefficients de parenté de la table diallèle.

Relation	$\phi$	$\phi'$	$\psi$	$\psi'$
Parent-enfant	2	0	1	0
Frères	1	1	$\frac{1}{6}$	$\frac{1}{6}$
Demi-Frères	1	0	$\frac{1}{6}$	0

On obtient les trois expressions des covariances :

$$\text{Cov(PO)} = \frac{1}{2} \sigma_A^2 + \frac{1}{6} \sigma_D^2 + \frac{1}{4} \sigma_{AA}^2 + \frac{1}{12} \sigma_{AD}^2 + \frac{1}{36} \sigma_{DD}^2 + \dots\dots\dots$$

$$\begin{aligned} \text{Cov(FS)} = & \frac{1}{2} \sigma_A^2 + \frac{2}{9} \sigma_D^2 + \frac{1}{12} \sigma_T^2 + \frac{1}{36} \sigma_Q^2 + \frac{1}{4} \sigma_{AA}^2 + \frac{1}{9} \sigma_{AD}^2 + \frac{1}{24} \sigma_{AT}^2 + \frac{1}{72} \sigma_{AQ}^2 + \frac{4}{51} \sigma_{DD}^2 \\ & + \frac{1}{54} \sigma_{DT}^2 + \frac{1}{162} \sigma_{DQ}^2 + \frac{1}{144} \sigma_{TT}^2 + \frac{1}{432} \sigma_{TQ}^2 + \frac{1}{1296} \sigma_{QQ}^2 + \dots\dots\dots \end{aligned}$$

$$\text{Cov(HS)} = \frac{1}{4} \sigma_A^2 + \frac{1}{36} \sigma_D^2 + \frac{1}{16} \sigma_{AA}^2 + \frac{1}{144} \sigma_{AD}^2 + \frac{1}{1296} \sigma_{DD}^2 + \dots\dots\dots$$

#### F - Etude d'une population tétraploïde après une génération d'autofécondation

Soit F le coefficient de consanguinité de la population initiale, de moyenne  $\mu$  et  $P_0^O, P_1^O, P_2^O, P_3^O, P_4^O$  les fréquences des cinq structures génotypiques, monogénique, digénique (simplex et duplex), trigénique, tétragénique.

Après une génération d'autofécondation le coefficient de consanguinité  $F_1$  est :  $(1 - F_0) \frac{5}{6} = (1 - F_1)$  et les fréquences des cinq structures :

$P_0^1, P_1^1, P_2^1, P_3^1, P_4^1$  sont donnés par la relation matricielle :

$$\begin{bmatrix} p_0^1 \\ p_1^1 \\ p_2^1 \\ p_3^1 \\ p_4^1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{4} & \frac{1}{18} & \frac{1}{36} & 0 \\ 0 & \frac{1}{2} & \frac{4}{9} & \frac{2}{9} & 0 \\ 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & \frac{1}{6} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{4}{6} \\ 0 & 0 & 0 & 0 & \frac{1}{6} \end{bmatrix} \times \begin{bmatrix} p_0^0 \\ p_1^0 \\ p_2^0 \\ p_3^0 \\ p_4^0 \end{bmatrix}$$

Supposons que  $F_0 = 0$  alors  $F_1 = \frac{1}{6}$

$$p_0^0 = p_1^0 = p_2^0 = p_3^0 = 0 \quad p_4^0 = 1$$

ce qui entraîne

$$p_0^1 = p_1^1 = 0$$

$$p_2^1 = \frac{1}{6} \quad p_3^1 = \frac{1}{3} \quad p_4^1 = \frac{1}{6}$$

seules les structures digéniques (duplex), trigénique, et tétragénique sont représentées avec des fréquences  $\frac{1}{6}$ ,  $\frac{2}{3}$  et  $\frac{1}{6}$  respectivement.

### 1°) Moyenne de la population

#### a) Cas d'un locus

En utilisant le même modèle que précédemment on peut décomposer la valeur génotypique des différentes structures.

$$y_{iijj} = \mu + 2\alpha_i + 2\alpha_j + 4\beta_{ij} + \beta_{ii} + \beta_{jj} + 2\gamma_{iij} + 2\gamma_{jji} + \delta_{iijj}$$

$$E\{y_{iijj}\} = \mu + 2E\{\beta_{ii}\} + E\{\delta_{iijj}\}$$

En revenant aux définitions des paramètres on voit que  $E\{\beta_{ii}\}$  et  $E\{\delta_{iijj}\}$  ne sont pas nuls, en effet :

$$\beta_{ii} = v_i^2 \left( \sum_s p_s \Lambda_s \right)^2$$

$$E\{\beta_{ii}\} = \sum_i p_i v_i^2 \left( \sum_s p_s \Lambda_s \right)^2 \quad \text{qui n'est pas nul}$$

mais

$$\begin{aligned} E\{\gamma_{ij}\} &= \sum_{ij} p_i p_j v_i^2 v_j \left( \sum_s p_s \Lambda_s \right) \\ &= \sum_j p_j v_j \sum_i p_i v_i^2 \left( \sum_s p_s \Lambda_s \right) = 0 \end{aligned}$$

$$\text{car} \quad \sum_j p_j v_j = 0$$

$$y_{iijk} = \mu + 2\alpha_i + \alpha_j + \alpha_k + 2\beta_{ij} + 2\beta_{ik} + \beta_{jk} + \beta_{ii} + \gamma_{ijj} + \gamma_{iik} + 2\gamma_{ijk} + \delta_{iijk}$$

$$E\{y_{iijk}\} = \mu + E\{\beta_{ii}\}$$

$$\begin{aligned} y_{ijk1} &= \mu + \alpha_i + \alpha_j + \alpha_k + \alpha_1 + \beta_{ij} + \beta_{ik} + \beta_{il} + \beta_{jk} + \beta_{jl} + \beta_{kl} + \gamma_{ijk} + \gamma_{ijl} + \gamma_{ikl} \\ &\quad + \gamma_{jkl} + \delta_{ijk1} \end{aligned}$$

$$E\{y_{ijk1}\} = \mu$$

La moyenne de la population est donc :

$$\mu_I = \frac{1}{6} E\{y_{iijj}\} + \frac{2}{3} E\{y_{iijk}\} + \frac{1}{6} E\{y_{ijk1}\}$$

$$= \frac{1}{6} (\mu + 2E\{\beta_{ii}\} + E\{\delta_{iijj}\}) + \frac{2}{3} (\mu + E\{\beta_{ii}\}) + \frac{1}{6} \mu$$

$$\mu_I = \mu + E\{\beta_{ii}\} + \frac{1}{6} E\{\delta_{iijj}\}$$

b) Cas de deux loci

Les calculs se font simplement, en supposant que les paramètres  $\beta_{ii}$  et  $\delta_{ii jj}$  ont même moyenne pour les deux loci.

$$E\{y_{ii jj, 11 mm}\} = \mu + 4E\{\beta_{ii}\} + 2E\{\delta_{ii jj}\} + 4E\{(\beta\beta)_{ii, 11}\} + 4E\{(\beta\delta)_{ii, 11 mm}\} \\ + E\{(\delta\delta)_{ii jj, 11 mm}\}$$

La décomposition de  $y_{ii jj, 11 mm}$  est facile mais longue on notera seulement le résultat :

$$E\{y_{ii jj, 11 mm}\} = \mu + 2E\{\beta_{ii}\} + E\{(\beta\beta)_{ii, 11}\}$$

Pour la structure tétragénique il est évident que l'espérance mathématique de la valeur génotypique est  $\mu$ .

$$\mu_I = \frac{1}{6} (\mu + 4E\{\beta_{ii}\} + 2E\{\delta_{ii jj}\} + 4E\{(\beta\beta)_{ii, 11}\} + 4E\{(\beta\delta)_{ii, 11 mm}\} + E\{(\delta\delta)_{ii jj, 11 mm}\}) \\ + \frac{2}{3} (\mu + 2E\{\beta_{ii}\} + E\{(\beta\beta)_{ii, 11}\}) + \frac{1}{6} \mu \\ \mu_I = \mu + 2E\{\beta_{ii}\} + \frac{1}{3} E\{\delta_{ii jj}\} + \frac{4}{3} E\{(\beta\beta)_{ii, 11}\} + \frac{2}{3} E\{(\beta\delta)_{ii, 11 mm}\} \\ + \frac{1}{6} E\{(\delta\delta)_{ii jj, 11 mm}\}$$

2°) Variance de la population

a) Cas d'un locus

$$\sigma_I^2 = \frac{1}{6} E\{y_{ii jj}^2\} + \frac{2}{3} E\{y_{ii jj}^2\} + \frac{1}{6} E\{y_{ij kl}^2\} - (\mu_I)^2 \\ = \frac{1}{6} E\{(y_{ii jj} - \mu)^2\} + \frac{2}{3} E\{(y_{ii jj} - \mu)^2\} + \frac{1}{6} E\{(y_{ij kl} - \mu)^2\} \\ - (\mu_I - \mu)^2$$

Et en remarquant que  $E\{(y_{ijk1} - \mu)^2\} = \sigma_G^2$

$$\sigma_I^2 = \frac{1}{6} \sigma_G^2 + \frac{1}{6} E\{(y_{iijj} - \mu)^2\} + \frac{2}{3} E\{(y_{iijk} - \mu)^2\} - (\mu_I - \mu)^2$$

$$\sigma_G^2 = \sigma_A^2 + \sigma_D^2 + \sigma_T^2 + \sigma_Q^2$$

Le calcul des différents termes de  $\sigma_I^2$  est long mais ne présente pas de difficultés. Les résultats sont les suivants :

$$y_{iijj} - \mu = 2\alpha_i + 2\alpha_j + 4\beta_{ij} + \beta_{ii} + \beta_{jj} + 2\gamma_{iij} + 2\gamma_{jji} + \delta_{iijj}$$

$$\begin{aligned} E\{(y_{iijj} - \mu)^2\} &= 8E(\alpha_i^2) + 16E(\beta_{ij}^2) + 2E(\beta_{ii}^2) + 8E(\gamma_{iij}^2) + E(\delta_{iijj}^2) \\ &\quad + 8E(\alpha_i \beta_{ii}) + 16E(\alpha_i \gamma_{jji}) + 8E(\alpha_i \delta_{iijj}) + 32E(\beta_{ij} \gamma_{iij}) \\ &\quad + 8E(\beta_{ij} \delta_{iijj}) + 8E(\beta_{ii} \gamma_{jji}) + 4E(\beta_{ii} \delta_{iijj}) \\ &\quad + 8E(\gamma_{iij} \gamma_{jji}) + 8E(\gamma_{iij} \delta_{iijj}) + 2E(\beta_{ii} \beta_{jj}) \end{aligned}$$

$$\begin{aligned} y_{iijk} - \mu &= 2\alpha_i + \alpha_j + \alpha_k + 2\beta_{ij} + 2\beta_{ik} + \beta_{jk} + \beta_{ii} + \gamma_{iij} + \gamma_{iik} + 2\gamma_{ijk} \\ &\quad + \delta_{iijk} \end{aligned}$$

$$\begin{aligned} E\{(y_{iijk} - \mu)^2\} &= 6E(\alpha_i^2) + 9E(\beta_{ij}^2) + 2E(\gamma_{iij}^2) + 4E(\gamma_{ijk}^2) + E(\delta_{iijk}^2) \\ &\quad + 2E(\alpha_i \beta_{ii}) + 4E(\alpha_j \gamma_{iij}) + 8E(\beta_{ij} \gamma_{iij}) + 2E(\beta_{jk} \delta_{iijk}) \\ &\quad + 4E(\gamma_{ijk} \delta_{iijk}) \end{aligned}$$

D'où en réunissant les différents termes, l'expression de la variance génétique de la population des descendants d'autofécondation :

$$\begin{aligned}\sigma_I^2 = & \frac{1}{6} \sigma_G^2 - (\mu_I - \mu)^2 + \left(\frac{8}{6} + \frac{12}{3}\right) E(\alpha_i^2) + \left(\frac{16}{6} + \frac{18}{3}\right) E(\beta_{ij}^2) + \frac{8}{3} E(\gamma_{ijk}^2) \\ & + \left(\frac{2}{6} + \frac{2}{3}\right) E(\beta_{ii}^2) + \left(\frac{8}{6} + \frac{4}{3}\right) E(\gamma_{iij}^2) + \frac{1}{6} E(\delta_{iijj}^2) + \frac{2}{3} E(\delta_{iijk}^2) \\ & + \left(\frac{4}{3} + \frac{16}{6}\right) E(\alpha_i \beta_{ii}) + \left(\frac{8}{3} + \frac{16}{6}\right) E(\alpha_i \gamma_{jji}) + \frac{8}{6} E(\alpha_i \delta_{iijj}) \\ & + \left(\frac{32}{6} + \frac{16}{3}\right) E(\beta_{ij} \gamma_{iij}) + \frac{8}{6} E(\beta_{ij} \delta_{iijj}) + \frac{4}{3} E(\beta_{jk} \delta_{iijk}) + \frac{8}{6} E(\beta_{ii} \gamma_{jji}) \\ & + \frac{4}{6} E(\beta_{ii} \delta_{iijj}) + \frac{8}{6} E(\gamma_{iij} \gamma_{jji}) + \frac{8}{6} E(\gamma_{iij} \delta_{iijj}) + \frac{8}{3} E(\gamma_{ijk} \delta_{iijk}) \\ & + \frac{2}{6} \left[ E(\beta_{ii}) \right]^2\end{aligned}$$

et après simplification :

$$\begin{aligned}\sigma_I^2 = & \frac{3}{2} \sigma_A^2 + \frac{29}{18} \sigma_D^2 + \frac{5}{6} \sigma_T^2 + \frac{1}{6} \sigma_Q^2 \\ & + E(\beta_{ii}^2) + \frac{8}{3} E(\gamma_{iij}^2) + \frac{1}{6} E(\delta_{iijj}^2) + \frac{2}{3} E(\delta_{iijk}^2) + 4E(\alpha_i \beta_{ii}) + \frac{16}{3} E(\alpha_j \gamma_{iij}) \\ & + \frac{4}{3} E(\alpha_i \delta_{iijj}) + \frac{32}{3} E(\beta_{ij} \gamma_{iij}) + \frac{4}{3} E(\beta_{ij} \delta_{iijj}) + \frac{4}{3} E(\beta_{jk} \delta_{iijk}) + \frac{4}{3} E(\beta_{ii} \gamma_{jji}) \\ & + \frac{2}{3} E(\beta_{ii} \delta_{iijj}) + \frac{4}{3} E(\gamma_{iij} \gamma_{jji}) + \frac{4}{3} E(\gamma_{iij} \delta_{iijj}) + \frac{8}{3} E(\gamma_{ijk} \delta_{iijk}) - \frac{2}{3} \left[ E(\beta_{ii}) \right]^2 \\ & - \frac{1}{36} \left[ E(\delta_{iijj}) \right]^2 - \frac{1}{3} E(\beta_{ii}) E(\delta_{iijj})\end{aligned}$$

b) Cas de deux loci

$$\sigma_I^2 = \frac{1}{6} \sigma_G^2 - (\mu_I - \mu)^2 + \frac{1}{6} E\{(y_{iijj,11mm})^2\} + \frac{2}{3} E\{(y_{iijk,11mm})^2\} \text{ avec}$$

$$\sigma_G^2 = \sigma_A^2 + \sigma_D^2 + \sigma_T^2 + \sigma_Q^2 + \sigma_{AA}^2 + \sigma_{AD}^2 + \sigma_{AT}^2 + \sigma_{AQ}^2 + \sigma_{DD}^2 + \sigma_{DT}^2 + \sigma_{DQ}^2 + \sigma_{TT}^2 + \sigma_{TQ}^2 + \sigma_{QQ}^2$$

en utilisant la même méthode que précédemment qui est simple mais longue on aboutit au résultat suivant en ne tenant compte que de termes entrant dans la décomposition de la variance génétique (voir annexe 10 pour les autres termes).

$$\sigma_I^2 = \frac{3}{2} \sigma^2 + \frac{25}{18} \sigma_D^2 + \frac{5}{6} \sigma_T^2 + \frac{1}{6} \sigma_Q^2 + \frac{7}{3} \sigma_{AA}^2 + \frac{23}{9} \sigma_{AD}^2 + \frac{7}{6} \sigma_{AT}^2 + \frac{1}{6} \sigma_{AQ}^2 + \frac{77}{27} \sigma_{DD}^2$$

$$+ \frac{25}{12} \sigma_{DT}^2 + \frac{1}{6} \sigma_{DQ}^2 + \frac{5}{6} \sigma_{TT}^2 + \frac{1}{6} \sigma_{TQ}^2 + \frac{1}{6} \sigma_{QQ}^2 + \dots$$

### 3°) Covariances entre apparentés

#### a) Cas d'un locus

#### \* Covariance entre frères : $\text{cov}(\text{FS})_I$

C'est la covariance entre les mesures de deux descendants d'un même individu autofécondé.

Le résultat est le suivant :

$$\begin{aligned} \text{Cov}(\text{FS})_I = & \sigma_A^2 + \frac{25}{36} \sigma_D^2 + \frac{1}{4} \sigma_T^2 + \frac{1}{36} \sigma_Q^2 \\ & + 2E(\alpha_i \beta_{ii}) + \frac{35}{9} E(\alpha_i \gamma_{ijj}) + \frac{2}{3} E(\alpha_i \delta_{iijj}) + \frac{1}{4} E(\beta_{ii}^2) \\ & + \frac{3}{4} E(\beta_{ii} \beta_{jj}) + \frac{10}{3} E(\beta_{ij} \gamma_{iij}) + E(\beta_{ii} \gamma_{jji}) + \frac{5}{18} E(\beta_{ij} \delta_{iijj}) \\ & + \frac{10}{9} E(\beta_{jk} \delta_{iijk}) + \frac{5}{27} E(\beta_{ii} \delta_{iijj}) + \frac{1}{3} E(\gamma_{iij}^2) + \frac{1}{3} E(\gamma_{iij} \gamma_{jji}) \\ & + \frac{2}{3} E(\gamma_{iij} \gamma_{kkj}) + \frac{1}{3} E(\gamma_{ijk} \delta_{iijk}) + \frac{1}{9} E(\gamma_{iij} \delta_{iijj}) \\ & + \frac{4}{27} E(\gamma_{ijj} \delta_{lkkij}) + \frac{1}{216} E(\delta_{iijj}^2) + \frac{1}{27} E(\delta_{iijk}^2) + \frac{2}{27} E(\delta_{iijk} \delta_{jjik}) \\ & + \frac{1}{27} E(\delta_{iijj} \delta_{ijkk}) + \frac{1}{216} E(\delta_{iijj} \delta_{kkll}) - \left[ E(\beta_{ii}) + \frac{1}{6} E(\delta_{iijj}) \right]^2 \end{aligned}$$



**\*\* Covariance entre demi-frères :  $\text{cov}(\text{HS})_I$**

C'est la covariance entre la mesure d'un descendant d'un individu I autofécondé et la mesure d'un descendant du croisement de I par un autre individu.

Le résultat est le suivant :

$$\begin{aligned} \text{Cov}(\text{HS})_I = & \frac{1}{6} \sigma_A^2 + \frac{1}{36} \sigma_D^2 + \frac{1}{2} E(\alpha_i \beta_{ii}) + E(\alpha_i \gamma_{ijj}) + \frac{1}{3} E(\beta_{ij} \gamma_{iij}) \\ & + \frac{1}{36} E(\beta_{ij} \delta_{iijj}) + \frac{1}{9} E(\beta_{jk} \delta_{iijj}) - \left[ E(\beta_{ii}) + \frac{1}{6} E(\delta_{iijj}) \right]^2 \end{aligned}$$

**\*\*\* Covariance parent-enfant :  $\text{cov}(\text{PO})_I$**

C'est la covariance entre les mesures d'un individu et l'un de ces descendants par autofécondation.

Le résultat est le suivant :

$$\begin{aligned} \text{Cov}(\text{PO})_I = & \sigma_A^2 + \frac{5}{6} \sigma_D^2 + \frac{1}{2} \sigma_T^2 + \frac{1}{6} \sigma_Q^2 + E(\alpha_i \beta_{ii}) + 2E(\alpha_i \gamma_{ijj}) + \frac{1}{3} E(\alpha_i \delta_{iijj}) \\ & + 2E(\beta_{ij} \gamma_{iij}) + \frac{1}{6} E(\beta_{ij} \delta_{iijj}) + \frac{2}{3} E(\beta_{jk} \delta_{iijj}) + \frac{2}{3} E(\gamma_{ijk} \delta_{iijj}) \end{aligned}$$

**b) Cas de deux loci**

Ici seuls les termes entrant dans la décomposition de la variance génétique ont été calculés.

**\* Covariance entre frères :  $\text{cov}(\text{FS})_I$**

$$\begin{aligned} \text{Cov}(\text{FS})_I = & \sigma_A^2 + \frac{25}{36} \sigma_D^2 + \frac{1}{4} \sigma_T^2 + \frac{1}{36} \sigma_Q^2 + \frac{13}{12} \sigma_{AA}^2 + \frac{29}{36} \sigma_{AD}^2 + \frac{31}{108} \sigma_{AT}^2 \\ & + \frac{1}{36} \sigma_{AQ}^2 + \frac{667}{972} \sigma_{DD}^2 + \frac{1}{4} \sigma_{DT}^2 + \frac{1}{36} \sigma_{DQ}^2 + \frac{7}{36} \sigma_{TT}^2 + \frac{1}{36} \sigma_{TQ}^2 + \frac{1}{36} \sigma_{QQ}^2 \\ & + \dots \end{aligned}$$

**\*\*** Covariance entre demi-frères :  $\text{cov}(\text{HS})_{\text{I}}$

$$\text{Cov}(\text{HS})_{\text{I}} = \frac{1}{6} \sigma_{\text{A}}^2 + \frac{1}{36} \sigma_{\text{D}}^2 + \frac{7}{24} \sigma_{\text{AA}}^2 + \frac{7}{72} \sigma_{\text{AD}}^2 + \frac{29}{648} \sigma_{\text{DD}}^2 + \dots$$

**\*\*\*** Covariance parent-enfant :  $\text{cov}(\text{PO})_{\text{I}}$

$$\begin{aligned} \text{Cov}(\text{PO})_{\text{I}} = & \sigma_{\text{A}}^2 + \frac{5}{6} \sigma_{\text{D}}^2 + \frac{1}{2} \sigma_{\text{T}}^2 + \frac{1}{6} \sigma_{\text{Q}}^2 + \sigma_{\text{AA}}^2 + \frac{5}{6} \sigma_{\text{AD}}^2 + \frac{1}{2} \sigma_{\text{AT}}^2 + \frac{1}{6} \sigma_{\text{AQ}}^2 \\ & + \frac{1}{3} \sigma_{\text{DD}}^2 + \frac{4}{9} \sigma_{\text{DT}}^2 + \frac{1}{6} \sigma_{\text{DQ}}^2 + \frac{1}{3} \sigma_{\text{TT}}^2 + \frac{1}{6} \sigma_{\text{TQ}}^2 + \frac{1}{6} \sigma_{\text{QQ}}^2 + \dots \end{aligned}$$

Nous avons étudié la décomposition des différentes covariances entre apparentés dans les populations des hybrides, des parents et des descendants d'autofécondations, une tentative d'utilisation de ces résultats, joints aux tableaux d'analyse de variance, va être l'objet de la seconde partie.

DEUXIEME PARTIE : Application à l'étude d'une population allogame autotétraploïde :

La luzerne du Puits.

I - PRINCIPES

Les considérations théoriques développées dans la première partie vont nous permettre d'étudier dans un premier temps les aptitudes à la combinaison, les effets réciproques et les effets maternels d'un échantillon aléatoire d'une population de plantes allogames autotétraploïdes.

Dans un second temps, moyennant un certain nombre d'hypothèses restrictives, nous pourrions essayer d'apprécier la variabilité des effets géniques.

A) L'étude des aptitudes à la combinaison des effets réciproques et effets maternels se font à l'aide des tableaux II, III, IV, V (Ière PARTIE).

B) Appréciation de la variabilité des effets géniques.

Nous avons obtenu la décomposition de n variances et covariances en fonction des composants de la variance génétique :

$$\sigma_A^2, \sigma_D^2, \sigma_T^2, \sigma_Q^2, \sigma_{AA}^2, \sigma_{AD}^2 \dots\dots\dots$$

Nous pouvons écrire :

$$Z_i = a_i \sigma_A^2 + d_i \sigma_D^2 + t_i \sigma_T^2 + q_i \sigma_Q^2 + (aa)_i \sigma_{AA}^2 + \dots\dots\dots$$

Nous avons vu que, pour les  $Z_i$  calculés à partir des autofécondations, d'autres termes entrent dans la décomposition. Ce sont les termes qui tiennent compte des interactions entre allèles identiques; nous allons supposer que ces termes interviennent peu. En effet les interactions entre allèles identiques peuvent en première approximation être supposés négli-

geables ce qui biologiquement peut se concevoir; cette hypothèse (I) peut évidemment être contestée, il est possible que les résultats nous conduisent à refuser cette éventualité.

Nous allons, en faisant l'hypothèse (I), estimer dans un premier temps les quatre composants d'additivité et de dominance à l'aide des quatre premières équations. Ensuite, à l'aide des huit équations, nous essaierons d'estimer en plus les deux premières composantes de l'épistasie ( $\sigma_{AA}^2$  et  $\sigma_{AD}^2$ ) qui d'après DUDLEY (1963) semblent être les plus importantes.

Dans ces deux cas, pour obtenir des composants de la variance génétique, la méthode sera la suivante :

Si nous nous trouvons en présence de l'estimation de n composants pour n  $Z_i$ , nous écrirons :

$$Z_i = \sum_{j=1}^n a_{ij} \sigma_i^2$$

La résolution du système de n équations à n inconnues donnera des estimations des  $\sigma_i^2$ .

Si nous nous trouvons en présence de l'estimation de p composants pour n  $Z_i$  et  $p < n$ , nous écrirons :

$$Z_i = \sum_{j=1}^p a_{ij} \sigma_i^2 + \epsilon_i$$

Nous chercherons les  $\sigma_i^2$  qui rendent minimum les  $\epsilon_i$  en utilisant la méthode des moindres carrés :

soit  $Z$  le vecteur colonne ayant pour composants les  $Z_i$ ; (variables dépendantes);  
 $A$  la matrice des  $a_{ij}$  (variables indépendantes)  
 $\sigma$  le vecteur colonne ayant pour composants les  $\sigma_i$ ;  
 $\epsilon$  le vecteur colonne ayant pour composants les  $\epsilon_i$ ;  
 Nous pouvons écrire :

$$Z = A \sigma + \epsilon$$

et la solution matricielle des équations normales est :

$$G = (H' H)^{-1} (H' Z)$$

## II - MATERIEL ET METHODES

L'expérience a été menée à la Station d'Amélioration des Plantes Fourragères de Lusignan (I.N.R.A.).

### 1°) Choix de la population

Le choix de la population a été fait en fonction des hypothèses définies dans le modèle génétique (Ière partie III-A). On peut considérer que les cinq hypothèses nécessaires aux calculs sont approximativement vérifiées pour le matériel végétal choisi : la luzerne du puits.

#### 1 - Ségrégation chromosomique :

La constitution de la luzerne est fortement autotétraploïde (DEMARLY, 1963). Ceci ne fait pour ainsi dire plus aucun doute. La ségrégation est chromosomique  $\alpha=0$  : il est peu vraisemblable que tous les gènes conditionnant le caractère soient situés dans une zone proche du centromère. On doit donc penser que dans certains cas,  $\alpha$  prend une valeur non nulle; quoiqu'il en soit, les modifications apportées aux résultats en faisant l'hypothèse de ségrégation chromatidique ne sont jamais très importantes.

#### 2 - Epistasie quelconque et absence de linkage :

Là encore, l'hypothèse est simplificatrice mais la luzerne ayant 32 chromosomes ( $n=8$ ) les probabilités de linkage sont relativement limitées.

#### 3 - Nombre d'allèles quelconques à chaque locus.

#### 4 - Parents consanguins ou non issus d'une population panmictique sans sélection à l'équilibre :

La population de luzerne du Puits étant en multiplication depuis plusieurs années,

on peut supposer qu'elle est en panmixie à l'équilibre, ou du moins que les déviations à la panmixie ne peuvent changer dans de grandes proportions la variance génétique.

D'autre part, la luzerne étant une plante allogame, le pourcentage d'autofécondations au cours des générations de multiplications ne peut entraîner de déviations à la panmixie plus importante que celle due au hasard.

On peut donc considérer la population en panmixie à l'équilibre.

#### 5 - Pas d'interaction génotype x environnement :

Les conditions de l'expérience sont telles que l'on peut considérer le milieu suffisamment homogène pour que l'interaction génotype x environnement soit négligeable.

#### 2°) L'obtention du matériel.

Un échantillon de 10 plantes a été tiré au hasard de la population du Puits de la pépinière du domaine du chêne. Les 90 croisements entre les 10 plantes et autofécondations ont été réalisés. La pollinisation a été faite à la main avec une lame de verre, en transportant le pollen d'une inflorescence à l'autre. Avec cette technique, on obtient 97 % de fécondations croisées en apportant un excès de pollen étranger sans qu'il soit nécessaire de pratiquer la castration. Les autofécondations sont faites simplement en déclanchant les fleurs avec la lame de verre.

#### 3°) Plan d'expérience

L'essai comporte 6 blocs complets randomisés, soit 6 répétitions de plantes par <sup>10</sup> traitement, il y a 114 traitements répartis comme suit : 90 croisements, 10 autofécondations, 10 parents bouturés, l'échantillon de graines et l'échantillon de boutures de la population d'origine. Les quatre derniers échantillons sont mis en place pour servir de référence. Chaque traitement est représenté par une ligne de 10 plantes. La culture sans sol en serre

(DEMARLY et CHESNEAUX - 1966) a permis de réduire au minimum la place prise par l'essai et de maintenir constantes les conditions de température, d'éclairage et de nutrition. L'essai a été installé dans des bacs en fibro-ciment remplis à partir du bas de lits successifs de cailloux, gravier, sable grossier et sable fin, de façon à permettre l'alimentation des plantes par sub-irrigation; un réservoir muni d'une pompe permet la montée d'une solution nutritive deux fois par jour, avec trop plein, contrôle automatique de niveau et retour dans le réservoir par gravité, après drainage à travers le sable. La source de lumière était constituée de lampes situées au-dessus de chaque bac, commandées automatiquement de façon à placer les plantes en jour continu. Les plantules ont été mises en place dans les bacs après prégermination en boîte de Pétri. L'essai est resté en place environ 6 semaines, la coupe a été faite plante par plante.

#### 4°) Caractères étudiés

La sélection se fait en général sur le caractère le plus important des plantes, c'est-à-dire la vigueur. Pour représenter le rendement, le poids de matière verte paraît le mieux adapté, la hauteur et le nombre de feuilles sont en corrélation avec le rendement, tout au moins au stade jeune où les mesures sont faites.

Les trois caractères choisis pour l'étude ont donc été le poids en vert, la hauteur et le nombre de feuilles des plantes.

### III - ANALYSE DES RESULTATS

#### A - Examen des tableaux

Le tableau ci-après contient les moyennes générales.

	Poids	Hauteur	Nb feuilles
Hybrides	2,422	13,25	6,19
Population	2,184	12,60	5,89
Auto-fécondations	1,773	11,11	5,86

Un test de comparaison des moyennes pour chaque caractère montre que la moyenne des hybrides ne diffère pas significativement de la moyenne de la population, mais elle est significativement supérieure à la moyenne des descendants d'autofécondations.

Ceci montre que la population est en équilibre en panmixie, et que l'effet "d'inbreeding" est marqué.

#### 1°) Etude des hybrides

a) En annexe (12) les tableaux 1, 2, 3, présentent pour les trois caractères la table diallèle des moyennes des six blocs.

Dans l'ensemble de l'étude, les trois caractères se comportent de la même façon. Nous allons donc examiner le poids; les conclusions restent pratiquement valables pour la hauteur et pour le nombre de feuilles.

b) La décomposition aptitude générale, aptitude spécifique à la combinaison, effets réciproques, montre que les trois variances  $\sigma_g^2$ ,  $\sigma_s^2$ ,  $\sigma_r^2$ , sont significativement différentes de zéro au risque de 1 % (tableau I). On peut donc calculer les estimations  $\hat{\sigma}_g^2$ ,  $\hat{\sigma}_s^2$ ,  $\hat{\sigma}_r^2$ , ainsi que les variances de ces estimations et les intervalles de confiance à 95 % (tableau II). Les variances ne sont là qu'à titre indicatif car les  $\hat{\sigma}^2$  ne sont pas



des variables gaussiennes et le calcul des intervalles de confiance à partir de ces variances n'est pas possible. SCHEFFE (1959) donne les formules pour calculer par approximation ces intervalles de confiance. Ce sont ces dernières qui ont été utilisées ici.

On peut remarquer que les limites de confiances sont relativement larges (0,02 à 0,3 pour l'estimation de la variance aptitude générale à la combinaison). L'utilisation de l'estimation ponctuelle ne peut donc être très précise, et se trouve de ce fait sujette à caution. Les limites de confiance de l'estimation de la variance  $\sigma_s^2$  sont plus proches l'une de l'autre que celles de  $\sigma_g^2$ . Cette meilleure précision tient au fait que le nombre de degrés de liberté de  $\sigma_s^2$  est supérieur à celui de  $\sigma_g^2$ .

c) Le tableau III présente la décomposition des effets réciproques. Les résultats semblent un peu surprenants à première vue. En effet, cette nouvelle décomposition change le sens donné aux paramètres  $g$  et  $s$ . Nous voyons que  $m_1$  effet maternel et  $r'_{ij}$  appelé effet réciproque résiduel prélèvent une grande partie des effets aptitude générale et aptitude spécifique à la combinaison. Ces derniers deviennent alors non significativement différents de zéro. L'interprétation de ces nouvelles variables est très difficile. Il est possible qu'en voulant dégager trop d'informations, on aboutisse à la chose suivante, les nouveaux effets de l'aptitude à la combinaison n'ont plus de sens.

On peut remarquer aussi que le test  $F$  de AGC est un test approximatif, ce qui peut être une cause d'imprécision. La présence d'un effet paternel, qui a déjà été remarquée dans d'autres études, mais qui est difficilement explicable, n'est pas exclue. On a voulu étendre le modèle mathématique, ce qui est toujours possible, mais l'interprétation biologique devient difficile.

TABLEAU I

ANALYSE DE LA VARIANCE SUR LES MOYENNES PARCELLAIRES

Caractère : poids en vert (dg) m = 2,422

Variation	d.d.l.	Somme des carrés	Carré moyen	F (d.d.l.)
Blocs	5	435	87	4,95 5(89-445) HS
Croisements	89	253	2,843	
Résiduelle	445	255	0,574	
Totale	529	943		

$$\sigma = 0,758$$

$$CV = 31,30 \%$$

Variation	d.d.l.	Somme des carrés	Carré moyen	F (d.d.l.)
Blocs	5	435		4,43 (9-35) HS 3,36 (35-445) HS 4,20 (45-445) HS
AGC	9	77	8,562	
ASC	35	68	1,931	
Effets récipro.	45	108	2,409	
Résiduelle	445	255	0,574	
Totale	539	943		

TABLEAU II

INTERVALLES DE CONFIANCE A 95 %

Caractère : poids en vert

	Estimation	Limite inférieure	Limite supérieure
$\hat{\sigma}_g^2$	0,06907	0,02000	0,27696
$\hat{\sigma}_s^2$	0,11308	0,05752	0,22665
$\hat{\sigma}_r^2$	0,15292	0,09346	0,27247

VARIANCES DES ESTIMATIONS

Caractère : poids en vert

	Estimation	Variance
$\hat{\sigma}_g^2$	0,06907	0,0017908
$\hat{\sigma}_s^2$	0,11308	0,0014899
$\hat{\sigma}_r^2$	0,15292	0,0018014

TABLEAU III

DECOMPOSITION DES EFFETS MATERNELS

Variation	d.d.l.	Somme des carrés	Carré moyen	F (d.d.l.)
Blocs	5	435		
AGC	9	77	8,562	0,65 NS
ASC	35	68	1,931	1,14 NS (35-36)
Effets mat.	9	47	5,222	3,08 HS (9-36)
Réciproques résiduelles	36	61	1,694	2,95 HS (36-445)
Résiduelle	445	255	0,57303	
Totale	539	943		

d) Le tableau IV montre la décomposition "entre" et "intra" parcelles. Le carré moyen intraparcelles nous permet de calculer la variance génétique de la population d'origine associée à une variance d'environnement individuelle,  $\sigma_G^2 + \sigma_f^2$ .

Le tableau V résume les résultats déduits des calculs précédents. La cov(HS) est faible, ce qui ne montre pas d'effet maternel très marqué et qui peut approuver l'utilisation des premiers calculs pour l'estimation de  $\sigma_g^2$ ,  $\sigma_s^2$ , et  $\sigma_r^2$ .

2°) Etude des clones

Le tableau VI montre l'analyse de variance pour la décomposition blocs, clones, ré-

TABLEAU IV

VARIANCE INTRAPARCELLAIRE

Variation	d.d.l.	Somme des carrés	Carré moyen
Entre-parcelles	539	8225	
Intra-parcelles	4644	7792	1,67786
Totale	5183	16017	

TABLEAU V

ESTIMATIONS DES VARIANCES ET COVARIANCES DEDUITES DE L'ETUDE DES HYBRIDES

Estimation	Poids	Hauteur	Nombre de feuilles
$\sigma_1^2$	1,67786	23,59290	1,21605
$\sigma_2^2$	0,80007	16,72268	0,85619
$\sigma_e^2$	0,57303	8,10112	0,33034
Cov(HS)	0,06907	2,72661	0,17880
Cov(FS)	0,32768	7,64985	0,41749
$\sigma_g^2 + \sigma_f^2$	2,08200	32,21446	1,74190

sidu, faite sur les totaux des deux coupes effectuées à un mois d'intervalle. Le test F montre que la variance entre clones est significativement différente de zéro. Représente-t-elle effectivement la variance génétique de la population ?

La variance génétique calculée à partir de graines et la variance entre clones sont du même ordre pour le poids et très différentes pour la hauteur et le nombre de feuilles. Ceci peut en partie s'expliquer par le fait que la quantité de matière fraîche synthétisée est sous contrôle génétique direct dans chaque génotype bouturé mais s'exprime différemment par des composantes biologiques telles que la hauteur et le nombre de feuilles très sensibles à la différenciation de l'élément bouturé .

Les variances intraparcellaires (tableau VI bis) semblent confirmer ces hypothèses.

### 3°) Etude des autofécondations.

Les deux tableaux VIII nous permettent (voir 1ère partie) de calculer des estimations de  $\sigma_I^2 + \frac{6}{7} \sigma_f^2$  et  $\text{cov}(\text{FS})_I - \frac{1}{7} \sigma_f^2$ .

Là encore, il n'est pas possible de séparer  $\sigma_I^2$  et  $\text{cov}(\text{FS})_I$  de  $\sigma_f^2$ , variance d'environnement individuel.

### 4°) Etude faisant intervenir hybrides, clones, autofécondations.

Ces calculs sont ceux des covariances parent-enfant, parent et sa descendance issue d'autofécondations, demi-frères (descendant d'autofécondations et hybride ayant un parent commun). Ils sont faits directement sur les moyennes. Les résultats sont rassemblés dans le tableau IX avec tous les précédents.

TABLEAU VI

ANALYSE DE LA VARIANCE ENTRE CLONES

Variation	d.d.l.	Somme des carrés	Carré moyen	F
Blocs	3	2,12835		
Clones	9	43,32781	4,70309	13,357 HS
Résiduelle	24	8,45055	0,35211	
Totale	36	52,90671		

$$\hat{\sigma} = 0,59339$$

$$m = 2,8995$$

$$CV = 20,52 \%$$

TABLEAU VI Bis

VARIANCE INTRAPARCELLE CLONES

	Poids	Hauteur	Nb de feuilles
1ère coupe	0,44262	65,12344	1,52188
2ème coupe	1,09386	84,34350	2,13640

TABLEAU VII

TABLEAU RECAPITULATIF DES CALCULS EFFECTUES SUR LES CLONES

caractère paramètre	Poids	Hauteur	nb feuilles
Moyenne des clous	1,4457	31,326	6,487
Estimation de la variance entre clones	0,271936	46,269750	0,542625

Le calcul est fait sur les moyennes de parcelles ayant en moyenne 8,17 plantes par parcelle (moyenne harmonique).



TABLEAU VIII

ANALYSE DES AUTOFECONDATIONS

Variation	d.d.l.	Somme des carrés	Carré moyen	F
Blocs	5	15,65046		
Plantes	9	17,83170	1,9813	4,38 S
Résiduelle	45	20,31663	0,45148	
Totale	59	53,79879		

moyenne = 1,7726

CV = 37,8 %

écart type = 0,6719

Variation	d.d.l.	Somme des carrés	Carré moyen
entre-parcelles	59	376,5882	
intra-parcelle	360	200,2695	0,5563
Totale	419	576,8577	

## B - Essai d'appréciation des composants de la variance génétique

Le tableau X résume les coefficients des différents composants de la variance génétique dans les décompositions des variances et covariances.

L'utilisation des quatre premières équations pour l'estimation de  $\sigma_A^2, \sigma_D^2, \sigma_T^2, \sigma_Q^2$ , donne les résultats rassemblés au tableau XI.

L'utilisation des 8 équations pour l'estimation de  $\sigma_A^2, \sigma_D^2, \sigma_T^2, \sigma_Q^2, \sigma_{AA}^2, \sigma_{AD}^2$ , donne les résultats rassemblés au tableau XII.

Nous allons examiner ce qui ressort de cette expérience. Si l'examen des chiffres ne permet pas de tirer des conclusions sur le plan biologique, en ce qui concerne cette étude particulière, le modèle mathématique demeure cependant parfaitement valable et peut être repris pour d'autres études dont le support expérimental permettrait de récolter des données plus fines.

La répétition de cette expérience sur plusieurs échantillons de cette même population, en contrôlant de plus en plus finement les causes d'incertitudes, notamment par une étude préalable du bouturage, permettrait de vérifier si les résultats sont constants ou non dans leur ensemble.

Dans le tableau XI, on peut penser que le fait d'avoir trouvé des composants négatifs, entraîne la présence d'épistasie. Mais la maintenance de ces composants négatifs, lorsqu'on ajoute les variances de l'épistasie (tableau XII), ayant semble-t-il plus d'importance par rapport aux autres, ne nous éclaire pas sur la répartition des différentes variations contenues dans la variance génétique.

Le fait de trouver des composants négatifs pourrait permettre de penser qu'il faut faire entrer d'autres termes dans la variance décomposition de la variance génétique. Mais

ceci implique une expérience beaucoup plus importante, qui ferait intervenir des individus ayant d'autres parentés que celles étudiées dans ce travail. Le nombre de relations de parenté qui entrent dans l'analyse diallèle est limité, si on ne tient pas compte des individus issus d'autofécondations, qui font intervenir des interactions entre allèles identiques (ou isoactifs) sur lesquelles on trouve peu de renseignements. D'autres types d'analyses analogues au diallèle ont été proposées : analyse triallèle, analyse "double cross hybrid". (RAWLINGS et COCKERHAM, 1961-1962).

TABLEAU IX

TABLEAU RECAPITULATIF DES ESTIMATIONS DES VARIANCES ET DES COVARIANCES

	Poids	Hauteurs	Nb de feuilles
Cov (HS)	0,06907	2,72661	0,17880
Cov (FS)	0,32768	7,64985	0,41749
$\sigma_G^2 + \sigma_f^2$	2,08200	32,21446	1,74190
Cov (PO)	0,09370	6,99287	0,18727
Cov (HS) <sub>I</sub>	0,08374	3,96032	0,24958
Cov (PO) <sub>I</sub>	0,15454	9,56303	0,29587
Cov (FS) <sub>I</sub> - $\frac{1}{7} \sigma_f^2$	0,17550	8,98104	0,42616
$\sigma_I^2 + \frac{6}{7} \sigma_f^2$	0,22186	7,41152	0,46670
$\sigma^2$ Clones	2,32175	378,0239	4,43320

8,17 moyenne harmonique des nombres de plantes par parcelles

TABLEAU X

	$\sigma_A^2$	$\sigma_D^2$	$\sigma_T^2$	$\sigma_Q^2$	$\sigma_{AA}^2$	$\sigma_{AD}^2$	$\sigma_{AT}^2$	$\sigma_{AQ}^2$	$\sigma_{DD}^2$	$\sigma_{DT}^2$	$\sigma_{DQ}^2$	$\sigma_{TT}^2$	$\sigma_{TQ}^2$	$\sigma_{QQ}^2$	
$\sigma_G^2$	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	.....
Cov(FS)	0,5000	0,2222	0,0833	0,0278	0,2500	0,1111	0,0417	0,0139	0,0494	0,0185	0,0062	0,0069	0,0023	0,0008	.....
Cov(HS)	0,2500	0,0278	0,0000	0,0000	0,0625	0,0069	0,0000	0,0000	0,0278	0,0000	0,0000	0,0000	0,0000	0,0000	.....
Cov(PO)	0,5000	0,1667	0,0000	0,0000	0,2500	0,0833	0,0000	0,0000	0,0008	0,0000	0,0000	0,0000	0,0000	0,0000	.....
$\sigma_I^2$	1,5000	1,6111	0,8333	0,1667	2,3333	2,5555	1,1667	0,1667	2,8518	2,4167	0,1667	0,8333	0,1667	0,1667	.....
Cov(FS) <sub>I</sub>	1,0000	0,6944	0,2500	0,0278	1,0833	0,8055	0,2870	0,0278	0,6862	0,2500	0,0278	0,1944	0,0278	0,0278	.....
Cov(HS) <sub>I</sub>	0,1667	0,0278	0,0000	0,0000	0,2917	0,0972	0,0000	0,0000	0,0447	0,0000	0,0000	0,0000	0,0000	0,0000	.....
Cov(PO) <sub>I</sub>	1,0000	0,8333	0,5000	0,1667	1,0000	0,8333	0,5000	0,1667	0,3333	0,4444	0,1667	0,3333	0,1667	0,1667	.....

TABLEAU XI

	Poids	Hauteur	Nb de feuilles
$\sigma_A^2$	0,32072	9,36678	0,88553
$\sigma_D^2$	- 0,39996	13,85685	- 1,53297
$\sigma_F^2$	3,53098	- 6,52662	4,48226
$\sigma_Q^2$	- 1,36974	15,51744	- 2,09292

TABLEAU XII

	Poids	Hauteur	Nb de feuilles
$\sigma_A^2$	0,517552	14,730580	0,909906
$\sigma_D^2$	- 0,333425	- 9,369738	- 1,598968
$\sigma_F^2$	- 1,467283	- 12,045534	- 0,445180
$\sigma_Q^2$	3,111294	36,797609	2,185345
$\sigma_{AA}^2$	- 0,162419	5,720519	0,353719
$\sigma_{AD}^2$	0,413970	- 3,601011	0,332967

## CONCLUSION

Dans le cas des plantes allogames, le but est d'amener un progrès génétique en même temps que d'aboutir à un matériel végétal présentant une homogénéité de comportement qui permette d'envisager des applications pratiques sur le plan agricole. Ces buts seront atteints d'autant plus rapidement que les génotypes seront mieux appréciés.

Le progrès de la sélection est conditionné par l'étendue de la variabilité génétique. L'étude de la répartition de cette variabilité entre additivité, dominance, épistasie, doit donc contribuer à aider les biologistes qui sont concernés par l'amélioration des plantes.

Les effets des gènes s'exprimant directement (additivité) ou par leurs interactions (dominance et épistasie) ont la possibilité de se transmettre d'une génération à la suivante. Les effets additifs, dont la variabilité est mesurée par la variance d'additivité sont essentiellement dus aux allèles eux-mêmes, alors que les effets de dominance et d'épistasie sont des propriétés de l'ensemble du génotype.

Il est bien évident que la contribution d'un individu au progrès de la sélection (si sa performance est supérieure à la moyenne de la population) ne sera pas de même nature selon qu'elle sera due à une plus grande additivité ou à un génotype particulier.

On voit donc l'intérêt pour la sélection de pouvoir dissocier ces différents types d'action des gènes et d'en connaître leur variabilité.

La taille "p" de l'échantillon nécessaire à la réalisation des croisements dialèles doit être selon la littérature supérieure à 10. DESSUREAUX (1960) pense qu'il faudrait prendre au moins 16 plantes pour obtenir des estimations valables. Le nombre de de-

grés de liberté de l'aptitude générale à la combinaison étant  $(p-1)$ , il faudrait donc que  $p$  soit le plus grand possible pour que la puissance de l'estimation soit bonne. Il va de soi que les difficultés de la réalisation des croisements de la table diallèle complète croissent également avec la taille de l'échantillon. Nous avons pris ici 10 plantes et l'obtention des graines n'a pas été simple. Nous sommes cependant à la limite indiquée par les différents auteurs consultés. La taille de l'échantillon un peu faible est une source de biais.

Une autre difficulté apparaît lorsqu'on augmente le nombre de plantes, celle du test d'un grand nombre de croisements dans des conditions homogènes. En effet, l'interaction génotype x milieu que l'on peut négliger lorsque la surface de l'essai n'est pas trop grande devient non négligeable si l'on veut faire intervenir un nombre important de croisements, les variances  $\sigma_e^2$ ,  $\sigma_f^2$ , en sont d'autant plus importantes.

L'estimation de la variance génétique est biaisée, qu'elle soit obtenue à partir de la variance intraparcellaire des hybrides, ou qu'elle soit obtenue à partir de la variance entre clones. En effet, dans le premier cas elle ne peut être dissociée de  $\sigma_f^2$ , dans le second l'interaction génotype x clonage est importante, surtout pour la hauteur des plantes.

Une autre source de biais de nature génétique est due au fait que la fécondation n'est pas au hasard dans la population mais est faite systématiquement sur l'échantillon tiré au hasard dans la population.

Enfin, les composants de l'épistasie d'ordre élevé ne sont peut-être pas négligeables, ce qui entraîne une imprécision dans l'estimation des premiers composants de la variance génétique.

Les difficultés rencontrées dans l'étude expérimentale seront certainement aplanies dans l'avenir. Le but de ces quelques pages est d'apporter une aide mathématique aux biologistes, qui soit susceptible de faire progresser la sélection et de la rendre plus efficace.



BIBLIOGRAPHIE

- BOUFFETTE (A.R.), BOUFFETTE (J.) - 1966.- Aspects probabilistes de la génétique des tétraploïdes. Université de Lyon. Séminaire de mathématiques appliquées.
- BOX (G.E.P.) - 1954.- Some theorems on quadratic forms applied in the study of variance problems. I. Ann. Math. Stat., vol. 25, pp. 290-302.
- BUSBICE (T.H.), WILSIE (C.P.) - 1966.- Inbreeding depression and heterosis in autotetraploids with application to Medicago Sativa. Euphytica, vol. 15, pp. 52-67.
- COCKERHAM (C.C.) - 1954.- An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. Genetica, vol. 39, n° , pp. 859-882.
- COCKERHAM (C.C.) - 1963.- Estimation of genetics variances. Statistical genetics and plant breeding. W.D. Hanson and H.F. Robinson, Washington, pp. 53-94.
- COMSTOCK (R.E.), ROBINSON (H.F.) - 1948.- The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance. Biometrics, vol. IV, n° 4, pp. 254-266.
- COMSTOCK (R.E.), ROBINSON (H.F.) - 1952.- Estimation of average dominance of genes. Heterosis. Gowen, Iowa, pp. 495-516.
- CRUMP (S.L.) - 1946.- The estimation of variance components in analysis of variance. Biometric Bulletin, vol. II, n° 1, pp. 7-11.
- DEMARLY (Y.) - 1963.- Génétique des tétraploïdes et amélioration des plantes. Ann. Amélior. Plantes, vol. 13, n° 4, pp. 304-400.
- DEMARLY (Y.), CHESNEAUX (M.T.) - 1966.- La culture de la luzerne en conditions artificielles. Ann. Amélior. Plantes, vol. 16, n° 3, pp. 299-305.

DESSUREAUX (L.) - 1959.- Introduction to the autotetraploid diallel.

Canad. J. Genet. Cytol., vol. I, pp. 94-101.

DESSUREAUX (L.) - 1960.- Analysis of continuous variation in autotetraploid.

Rep. 7th. Alfala Impr. Conf., pp. 16-18.

EISENHART (C.) - 1947.- The assumptions underlying the analysis of variance.

Biometrics, vol. III, n° 1, pp. 1-21.

FISHER (R.A.) - 1943.- Allowance for double reduction in the calculation of the genotypic frequencies with polysomic inheritance. Ann. Eugen. Lond., vol. 12, pp. 169-171.

GRIFFING (B.) - 1956 a.- A generalised treatment of the use of diallel crosses in quantitative inheritance. Heredity, vol. 10, n° 1, pp. 31-50.

GRIFFING (B.) - 1956 b.- Concept of general and specific combining ability in relation to diallel crossing systems. Aust. J. Biol. Sci. vol. 9, pp. 463-493.

GRIFFING (B.) - 1958.- Application of sampling variables in the identification of methods which yield unbiased estimates of genotypic variance components.

Aust. J. Biol. Sci., vol. 11, pp. 219-245.

HAYMAN (B.I.) - 1954 a.- The theory and analysis of diallel crosses.

Genetics, vol. 39, pp. 789-809.

HAYMAN (B.I.) - 1954 b.- The analysis of variance of diallel tables.

Biometrics, vol. X, pp. 235-244.

HAYMAN (B.I.) - 1957.- Interaction, heterosis and diallel crosses.

Genetics, vol. 42, pp. 336-355.

HAYMAN (B.I.) - 1958 a.- The theory and analysis of diallel crosses II.

Genetics, vol. 43, pp. 63-85.

- HAYMAN (B.I.) - 1958 b.- The separation of epistatic from additive and dominance variation in generation means. *Heredity*, vol. 12, pp. 371-390.
- HAYMAN (B.I.) - 1960.- The theory and analysis of diallel crosses III. *Genetics*, vol. 45, pp. 155-172.
- HAYMAN (B.I.), MATHER (K.) - 1955.- The description of genic interactions of continuous variation. *Biometrics*, vol. XI, pp. 69-82.
- HILL (J.) - 1954.- Effects of correlated gene distributions in the analysis of diallel crosses. *Heredity*, vol. XIX, n° 1, pp. 27-46.
- JINKS (J.L.) - 1954.- The analysis of continuous variation in a diallel cross of Nicotiana rustica varieties. *Genetics*, vol. 39, pp. 767-788.
- JINKS (J.L.) - 1955.- A survey of the genetical basis of heterosis in a variety of diallel crosses. *Heredity*, vol. IX, pp. 223-238.
- JINKS (J.L.) - 1956.- The  $F_2$  and backcross generations from a set of diallel crosses. *Heredity*, vol. X, pp. 1-30.
- JINKS (J.L.), BROADHURSTS (P.L.) - 1963.- Diallel analysis of litter size and body weight in rats. *Heredity*, vol. 18, pp. 319-336.
- JINKS (J.L.), HAYMAN (B.I.) - 1953.- The analysis of diallel crosses. *Maize Genetics Coop. New letters*, vol. 27, pp. 48-54.
- KEMPTHORNE (O.) - 1952.- Design and analysis experiment. John Wiley, New-York, 631 p.
- KEMPTHORNE (O.) - 1955 a.- The correlation between relatives in a simple autotetraploid population. *Genetics*, vol. 40, pp. 167-174.
- KEMPTHORNE (O.) - 1955 b.- The correlation between relatives in random mating populations. *Cold. Spr. Harb. Symp. Quant. Biol.*, vol. 20, pp. 60-75.

- KEMPTHORNE (O.) - 1955 c.- The theoretical values of correlations between relatives in random mating populations. Genetics, vol. 40, pp. 153-167.
- KEMPTHORNE (O.) - 1956.- The theory of the diallel crosses. Genetics, vol. 41, pp. 451-459.
- KEMPTHORNE (O.) - 1957.- An introduction to genetic statistics. John Wiley. New-York, 545 p.
- KEMPTHORNE (O.), CURNOW (O.) - 1961.- The partial diallel cross. Biometrics, vol. 17, pp. 229-250.
- LEVINGS (C.S.), DUDLEY (J.W.) - 1963.- Evaluation of certain matings designs for estimation of genetic variance in autotetraploid alfalfa. Crop Science, vol. 3, pp. 532-535.
- MALECOT (G.) - 1948.- Les mathématiques de l'hérédité. Masson et Cie, Paris, 61 p.
- RAWLING (J.O.), COCKERHAM (C.C.) - 1961.- Analysis of double cross hybrid populations. Biometrics, vol. 18, pp. 229-244.
- RAWLINGS (J.O.), COCKERHAM (C.C.) - 1962.- Trialallel analysis. Crop Sciences, pp. 229-229.
- REEVE (E.C.R.) - 1953.- Studies of quantitative inheritance and genetic correlation in progeny tests using different mating systems. J. Genet. vol. 51, pp. 520-542.
- SATTERTHWAITE ( . ) - 1946.- An approximate distribution of estimates of variance components. Biometric Bulletin, vol. 2, pp. 110-114.
- SCHEFFE (H.) - 1959.- The analysis of variance. John Wiley. New-York, 477 p.
- SPRAGUE (G.F.), TATUM (L.A.) - 1942.- General versus specific combining ability in single crosses of corn. J. Amer. Soc. Agron., vol. 34, pp. 923-932.
- WEARDEN (S.) - 1964.- Alternative analysis of diallel cross. Heredity, vol. 19, pp. 669-680.
- YATES (F.) - 1947.- Analysis of data from all possible reciprocal crosses between a set of parental lines. Heredity, vol. 1, pp. 287-301.

## A N N E X E S

## PLAN DES ANNEXES DE CALCUL

1. Estimation des paramètres du modèle I (groupe 3).
2. Calcul des espérances mathématiques des carrés moyens modèle I (Groupe 3).
3. Variances des estimations des paramètres modèle I (Groupe 3).
4. Calcul des espérances mathématiques des carrés moyens modèle II (groupe 3).
5. Calcul des variances des estimations des variances modèle II (groupe 3).
6. Expression de l'espérance mathématique d'un carré moyen.
7. Espérance mathématique du carré moyen intraparcellaire.
8. Variances et covariances des moyennes parcellaires.
9. Espérances des carrés moyens en fonction du modèle génétique.
10. Expression de la variance génétique d'une population tetraploïde après une génération d'autofécondation.
11. Formules approximatives des limites de confiance des estimations des variances.
12. Tableaux des moyennes.
13. Analyse de la variance sur les moyennes parcellaires.
14. Tableaux des intervalles de confiance des estimations des variances.
15. Tableaux des variances des estimations.
16. Décomposition des effets maternels.
17. Variances intraparcellaires.
18. Analyses de variance sur les clones.
19. Tableau des moyennes des clones.
20. Tableau des moyennes des autofécondations.
21. Tableaux des analyses de variance sur les autofécondations.

Estimations des paramètres du modèle I (Groupe 3).

$$y_{ijk} = \mu + g_i + g_j + s_{ij} + r_{ij} + b_k + l_{ijk},$$

$$i \neq j$$

$$s_{ij} = s_{ji} \quad r_{ij} = -r_{ji}$$

$$Y_{...} = \sum_{\substack{ijk \\ i \neq j}} y_{ijk}$$

$$= p(p-1)n\hat{\mu} + 2(p-1)n \sum_i \hat{g}_i + 2n \sum_{i < j} \hat{s}_{ij} + p(p-1) \sum_k \hat{b}_k$$

$$Y_{..k} = \sum_{i \neq j} y_{ijk}$$

$$= p(p-1)\hat{\mu} + 2(p-1) \sum_i \hat{g}_i + 2 \sum_{i < j} \hat{s}_{ij} + p(p-1)\hat{b}_k$$

$$Y_{i..} + Y_{.i.} = \sum_{\substack{jk \\ j \neq i}} (y_{ijk} + y_{jik})$$

$$= 2(p-1)n\hat{\mu} + 2(p-1)n\hat{g}_i + 2n \sum_{i \neq j} \hat{g}_j + 2n \sum_{j \neq i} \hat{s}_{ij} + 2(p-1) \sum_k \hat{b}_k$$

$$Y_{ij.} + Y_{ji.} = \sum_{\substack{k \\ i \neq j}} (y_{ijk} + y_{jik})$$

$$= 2n\hat{\mu} + 2n\hat{g}_i + 2n\hat{g}_j + 2n\hat{s}_{ij} + 2 \sum_k \hat{b}_k$$

$$Y_{ij.} - Y_{ji.} = \sum_{\substack{k \\ i \neq j}} (y_{ijk} - y_{jik})$$

$$= 2n\hat{r}_{ij}$$

Pour obtenir une solution on impose les conditions :

$$\sum_i \hat{g}_i = 0; \sum_j \hat{s}_{ij} = 0 \text{ pour tout } i; \sum_k \hat{b}_k = 0$$

Les solutions de ces équations sont les suivantes.

$$\hat{\mu} = \frac{Y_{...}}{p(p-1)n}$$

$$\hat{b}_k = \frac{Y_{..k}}{p(p-1)} - \frac{Y_{...}}{np(p-1)}$$

$$\hat{g}_i = \frac{Y_{i..} + Y_{.i.}}{2(p-2)n} - \frac{Y_{...}}{p(p-2)n}$$

$$\hat{s}_{ij} = \frac{Y_{ij.} + Y_{.ji}}{2n} - \frac{Y_{i..} + Y_{.i.} + Y_{j..} + Y_{.j.}}{2n(p-2)} + \frac{Y_{...}}{n(p-1)(p-2)}$$

$$\hat{r}_{ij} = \frac{Y_{ij.} - Y_{.ji}}{2n}$$



Annexe 2

Calcul des espérances mathématiques des carrés moyens  
modèle I (Groupe 3)

Les restrictions imposées aux paramètres sont les suivantes :

$$\sum_i q_i = 0, \sum_j \delta_j = 0 \text{ pour tout } i, \sum_k b_k = 0$$

$$Y_{...} = p(p-1)n\mu + 2(p-1)n \sum_i q_i + 2n \sum_{i < j} \delta_{ij} \\ + p(p-1) \sum_k b_k + \sum_{ijk} \ell_{ijk}$$

$$= p(p-1)n\mu + \sum_{ijk} \ell_{ijk}$$

$$E\left\{\frac{Y_{...}^2}{p(p-1)n}\right\} = p(p-1)n\mu^2 + \sigma^2$$

$$Y_{i..} + Y_{.i.} = 2(p-1)n\mu + 2(p-2)nq_i + \sum_{jk} (\ell_{ijk} + \ell_{jik})$$

$$E\left\{\frac{(Y_{i..} + Y_{.i.})^2}{2(p-2)n}\right\} = \frac{2p(p-1)n\mu^2}{p-2} + 2n(p-2) \sum_i q_i^2 + \frac{p(p-1)}{p-2} \sigma^2$$

$$E\{S_g\} = 2n(p-2) \sum_i q_i^2 + (p-1)\sigma^2$$

$$E\{CM_g\} = \sigma^2 + \frac{2n(p-2)}{p-1} \sum_i q_i^2$$

$$Y_{ij.} + Y_{ji.} = 2n\mu + 2nq_i + 2nq_j + 2n\delta_{ij} + 2 \sum_k b_k + \sum_k (\ell_{ijk} + \ell_{jik})$$

$$\sum_{i < j} (Y_{ij.} + Y_{ji.})^2 = 4n^2 \frac{p(p-1)}{2} \mu^2 + 4n^2 \sum_{i < j} (q_i + q_j)^2 + 4n^2 \sum_{i < j} \delta_{ij}^2 \\ + \sum_{i < j} \left( \sum_k \ell_{ijk} + \ell_{jik} \right)^2$$

$$E \left\{ \sum_{i < j} \frac{(Y_{ij} + Y_{ji})^2}{2n} \right\} = np(p-1)\mu^2 + 2n(p-2) \sum_i q_i^2 + 2n \sum_{i < j} \Delta_{ij}^2 + \frac{p(p-1)}{2} \sigma^2$$

$$\begin{aligned} E \{ S_A \} &= p(p-1)n\mu^2 + 2n(p-2) \sum_i q_i^2 + 2n \sum_{i < j} \Delta_{ij}^2 \\ &+ \frac{p(p-1)}{2} \sigma^2 - \frac{2p(p-1)^2 n \mu^2}{p-2} - 2n(p-2) \sum_i q_i^2 - \frac{p(p-1)}{p-2} \sigma^2 \\ &+ \frac{p^2(p-1)^2 n^2}{n(p-1)(p-2)} \mu^2 + \frac{p(p-1)n}{n(p-1)(p-2)} \sigma^2 \\ &= \frac{p(p-3)}{2} \sigma^2 + 2n \sum_{i < j} \Delta_{ij}^2 \end{aligned}$$

$$E \{ CM_A \} = \sigma^2 + 2n \frac{2}{p(p-3)} \sum_{i < j} \Delta_{ij}^2$$

$$Y_{ij} - Y_{ji} = 2n r_{ij} + \sum_k (l_{ijk} - l_{jik})$$

$$E \{ (Y_{ij} - Y_{ji})^2 \} = 4n^2 r_{ij}^2 + 2n \sigma^2$$

$$\sum_{i < j} \frac{E \{ (Y_{ij} - Y_{ji})^2 \}}{2n} = 2n \sum_{i < j} r_{ij}^2 + \frac{p(p-1)}{2} \sigma^2$$

$$E \{ CM_n \} = \sigma^2 + 2n \frac{2}{p(p-1)} \sum_{i < j} r_{ij}^2$$

# Annexe 3

Variance des estimations des paramètres modèle I (p. 3)

$$\text{var}(Y_{ijk}) = \hat{\sigma}^2 = \text{C.H. résiduel.}$$

$$\text{var}(Y_{ijk} - Y_{lmn}) = 2 \hat{\sigma}^2$$

$$\text{var}(\hat{\mu}) = \text{var}\left(\frac{\sum_{ijk} Y_{ijk}}{p(p-1)n}\right) = \frac{\hat{\sigma}^2}{p(p-1)n}$$

$$\hat{f}_i = \frac{Y_{i..} + Y_{.i.}}{2(p-2)n} - \frac{Y_{...}}{p(p-2)n}$$

$$\begin{aligned} \text{var}(\hat{f}_i) &= \frac{1}{4(p-2)^2 n^2} \left\{ \text{var}(Y_{i..}) + \text{var}(Y_{.i.}) + 2 \text{cov}(Y_{i..}, Y_{.i.}) \right\} \\ &\quad + \frac{1}{p^2(p-2)^2 n^2} \text{var}(Y_{...}) - \frac{2}{p(p-2)^2 n^2} \text{cov}(Y_{i..}, Y_{...}) \\ &\quad - \frac{1}{p(p-2)^2 n^2} \text{cov}(Y_{.i.}, Y_{...}) \end{aligned}$$

$$\text{cov}(Y_{i..}, Y_{.i.}) = \text{cov}\left(\sum_{\substack{j,k \\ j \neq i}} Y_{ijk}, \sum_{\substack{l,m \\ l \neq i}} Y_{lim}\right) = 0$$

$$\text{cov}(Y_{i..}, Y_{...}) = \sum_{\substack{j,k \\ j \neq i}} \text{var}(Y_{ijk}) = n(p-1)\hat{\sigma}^2$$

$$\begin{aligned} \text{var}(\hat{f}_i) &= \frac{1}{4(p-2)^2 n^2} \left\{ (p-1)n\hat{\sigma}^2 + (p-1)n\hat{\sigma}^2 + 0 \right\} \\ &\quad + \frac{1}{p^2(p-2)^2 n^2} \left\{ p(p-1)n\hat{\sigma}^2 \right\} - \frac{2}{p(p-2)^2 n^2} n(p-1)\hat{\sigma}^2 \end{aligned}$$

$$\text{var}(\hat{\gamma}_i) = \frac{p-1}{2n(p-2)} \hat{\sigma}^2$$

$$\begin{aligned} \hat{\delta}_{ij} &= \frac{Y_{ij.} + Y_{ji.}}{2n} - \frac{1}{2n(p-2)} (Y_{i..} + Y_{.i.} + Y_{j..} + Y_{.j.}) \\ &\quad + \frac{Y_{...}}{n(p-1)(p-2)} \end{aligned}$$

En tenant compte des relations suivantes :

$$\text{cov}(Y_{ij.}, Y_{ji.}) = 0$$

$$\text{cov}(Y_{i..}, Y_{j..}) = 0$$

$$\text{cov}(Y_{ij.}, Y_{i..}) = n \hat{\sigma}^2$$

$$\text{cov}(Y_{ij.}, Y_{.i.}) = 0$$

$$\text{cov}(Y_{ij.}, Y_{...}) = n \hat{\sigma}^2$$

On obtient :

$$\begin{aligned} \text{var}(\hat{\delta}_{ij}) &= \hat{\sigma}^2 \left[ \frac{1}{2n} + \frac{p-1}{n(p-2)^2} + \frac{p}{n(p-1)(p-2)^2} - \frac{2}{n(p-2)} \right. \\ &\quad \left. + \frac{2}{n(p-1)(p-2)} - \frac{4}{n(p-2)^2} + \frac{1}{n(p-2)^2} \right] \end{aligned}$$

$$\text{var}(\hat{\delta}_{ij}) = \frac{p-3}{2n(p-1)} \hat{\sigma}^2$$

$$\text{var}(\hat{q}_i - \hat{q}_j) = \frac{2(p-1)}{2np(p-2)} \hat{\sigma}^2 - 2 \text{cov}(\hat{q}_i, \hat{q}_j)$$

$$= \hat{\sigma}^2 \left[ \frac{1}{2n(p-2)^2} - \frac{2(p-1)}{np(p-2)^2} + \frac{p-1}{np(p-2)^2} \right]$$

$$\boxed{\text{var}(\hat{g}_i - \hat{g}_j) = \frac{\hat{\sigma}^2}{n(p-2)} \quad i \neq j}$$

$$\text{var}(\hat{\delta}_{ij} - \hat{\delta}_{ik}) = \text{var}(\hat{\delta}_{ij}) + \text{var}(\hat{\delta}_{ik}) - 2\text{cov}(\hat{\delta}_{ij}, \hat{\delta}_{ik})$$

$i \neq j \text{ and } j \neq k.$

$$= \hat{\sigma}^2 \left[ \frac{1}{n(p-2)} + \frac{2}{n(p-1)(p-2)} + \frac{p-1}{2n(p-2)^2} + \frac{3}{2n(p-2)^2} + \frac{p}{n(p-1)(p-2)^2} - \frac{4}{n(p-2)^2} \right]$$

$$\boxed{\text{var}(\hat{\delta}_{ij} - \hat{\delta}_{ik}) = \frac{p-3}{n(p-2)} \hat{\sigma}^2 \quad i \neq j, k, j \neq k}$$

Annexe 4.

Calcul des espérances mathématiques des carrés moyens  
modèle II (Groupe 3).

$$Y_{...} = p(p-1)n\mu + 2(p-1)n \sum_i g_i + 2n \sum_{i < j} \delta_{ij} + p(p-1) \sum_k b_k + \sum_{ijk} \epsilon_{ijk}$$

$$E(Y_{...}) = p^2(p-1)^2 n^2 \mu^2 + 4(p-1)^2 n^2 p \sigma_g^2 + 4n^2 \frac{p(p-1)}{2} \sigma_\delta^2 + p^2(p-1)^2 n \sigma_b^2 + p(p-1)n \sigma^2$$

$$Y_{i..} + Y_{.i.} = 2(p-1)n\mu - 2(p-1)n g_i + 2n \sum_{i \neq j} g_j + 2n \sum_{i \neq j} \delta_{ij} + 2(p-1) \sum_k b_k + \sum_{jk} (\epsilon_{ijk} + \epsilon_{jik})$$

$$\sum_i E \left\{ \frac{(Y_{i..} + Y_{.i.})^2}{2(p-2)n} \right\} = \frac{2(p-1)^2 n p}{p-2} \mu^2 + \frac{2p^2(p-1)n}{p-2} \sigma_g^2 + \frac{2np(p-1)}{p-2} \sigma_\delta^2 + \frac{p(p-1)\sigma^2}{p-2} + \frac{2p(p-1)^2}{p-2} \sigma_b^2$$

$$E \left\{ S_g \right\} = 2n(p-1)(p-2) \sigma_g^2 + 2n(p-1) \sigma_\delta^2 + (p-1) \sigma^2$$

$$E \left\{ CM_g \right\} = 2n(p-2) \sigma_g^2 + 2n \sigma_\delta^2 + \sigma^2$$

$$Y_{ij.} + Y_{ji.} = 2n\mu + 2n g_i + 2n g_j + 2n \delta_{ij} + 2 \sum_k b_k + \sum_k (\epsilon_{ijk} + \epsilon_{jik})$$

$$\sum_{i < j} E \left\{ \frac{(Y_{ij.} + Y_{ji.})^2}{2n} \right\} = p(p-1)n\mu^2 + 2p(p-1)n \sigma_g^2 + p(p-1)n \sigma_\delta^2 + p(p-1) \sigma_b^2 + \frac{p(p-1)}{2} \sigma^2$$

$$E \left\{ S_A \right\} = \frac{2np(p-3)}{2} \sigma_A^2 + \frac{p(p-3)}{2} \sigma^2$$

$$E \left\{ CM_A \right\} = 2n \sigma_A^2 + \sigma^2$$

$$Y_{ij.} - Y_{ji.} = 2n r_{ij} + \sum_k (l_{ijk} - l_{jik})$$

$$E \left\{ \frac{(Y_{ij.} - Y_{ji.})^2}{2n} \right\} = \frac{4n^2 \sigma_n^2}{2n} + \frac{2n \sigma^2}{2n}$$

$$\sum_{i < j} E \left\{ \frac{(Y_{ij.} - Y_{ji.})^2}{2n} \right\} = 2n \frac{p(p-1)}{2} \sigma_n^2 + \frac{p(p-1)}{2} \sigma^2$$

$$E \left\{ CM_n \right\} = 2n \sigma_n^2 + \sigma^2$$

## Annexe 5

Calcul des variances des estimations des variances  
modèle II (Groupe 3)

La formule suivante donne une approximation  
de la variance d'un carré moyen.

$$\text{var}(\text{CM}_0) = \frac{2}{\nu_0} (\text{CM}_0)^2$$

(Scheffe - 1959)

$\nu_0$  étant le nombre de degrés de liberté du carré  
moyen  $\text{CM}_0$

$$\hat{\sigma}_g^2 = \frac{1}{2r(p-2)} (\text{CM}_g - \text{CM}_0)$$

$$\text{var}(\hat{\sigma}_g^2) = \frac{1}{4r^2(p-2)^2} [\text{var}(\text{CM}_g) + \text{var}(\text{CM}_0)]$$

en faisant l'hypothèse que les carrés moyens sont  
indépendants.

$$\text{var}(\hat{\sigma}_g^2) = \frac{1}{4r^2(p-2)^2} \left[ \frac{2}{(p-1)} (\text{CM}_g)^2 + \frac{4}{p(p-3)} (\text{CM}_0)^2 \right]$$

$$\hat{\sigma}_0^2 = \frac{1}{2r} (\text{CM}_0 - \text{CM}_{\text{res}})$$



$$\text{var}(\hat{\sigma}_s^2) = \frac{1}{4n^2} \left[ \text{var}(M_s) + \text{var}(M_{res}) \right]$$

$$\text{var}(\hat{\sigma}_s^2) = \frac{1}{4n^2} \left[ \frac{4}{p(p-3)} (M_s)^2 + \frac{2}{v} (M_{res})^2 \right]$$

$$\sigma_r^2 = \frac{1}{2v} (M_r - M_{res})$$

$$\text{var}(\hat{\sigma}_r^2) = \frac{1}{4n^2} \left[ \text{var}(M_r) + \text{var}(M_{res}) \right]$$

$$\text{var}(\hat{\sigma}_r^2) = \frac{1}{4n^2} \left[ \frac{4}{p(p-1)} (M_r)^2 + \frac{2}{v} (M_{res})^2 \right]$$

Exercice 6

Exprimer la covariance mathématique d'un carré moyen

Démonstration de la formule suivante:

$$E \left\{ \sum_{l=1}^n \frac{(x_l - \bar{x})^2}{n-1} \right\} = \frac{1}{2} E \left\{ (x_l - x_{l'})^2 \right\} \quad l \neq l'$$

1°) Calcul de  $E \left\{ (x_l - x_{l'})^2 \right\}$ ;

$$\begin{aligned} E \left\{ (x_l - x_{l'})^2 \right\} &= E(x_l^2) + E(x_{l'}^2) - 2E(x_l x_{l'}) \\ &= 2E(x_l^2) - 2E(x_l x_{l'}) \end{aligned}$$

$$E(x_l^2) = \text{var}(x_l) + [E(x_l)]^2$$

$$\begin{aligned} E \left\{ (x_l - x_{l'})^2 \right\} &= 2 \text{var}(x_l) + 2[E(x_l)]^2 - 2E(x_l x_{l'}) \\ &= 2 \text{var}(x_l) - 2 \left[ E(x_l x_{l'}) - [E(x_l)]^2 \right] \\ &= 2 \text{var}(x_l) - 2 \left[ E(x_l x_{l'}) - E(x_l)E(x_{l'}) \right] \\ &= 2 \text{var}(x_l) - 2 \text{cov}(x_l, x_{l'}) \end{aligned}$$

$$\boxed{\frac{1}{2} E \left\{ (x_l - x_{l'})^2 \right\} = \text{var}(x_l) - \text{cov}(x_l, x_{l'})}$$

2°) Calcul de  $E \left\{ (x_e - \bar{x})^2 \right\}$

$$E \left\{ (x_e - \bar{x})^2 \right\} = E(x_e^2) - 2E(\bar{x}, x_e) + E(\bar{x}^2)$$

$$E(\bar{x}^2) = \text{var}(\bar{x}^2) + [E(\bar{x})]^2$$

$$= \frac{n \text{ var}(x_e) + n(n-1) \text{ cov}(x_e, x_{e'})}{n^2} + [E(\bar{x})]^2$$

$$= \frac{1}{n} \text{ var}(x_e) + \frac{n-1}{n} \text{ cov}(x_e, x_{e'}) + [E(x_e)]^2$$

$$E(\bar{x} x_e) = E\left(\frac{\sum_i x_i}{n}, x_e\right)$$

$$= \frac{n-1}{n} E(x_e x_{e'}) + \frac{1}{n} E(x_e^2)$$

$$\begin{aligned} E \left\{ (x_e - \bar{x})^2 \right\} &= \text{Var}(x_e) + [E(x_e)]^2 - \frac{2(n-1)}{n} E(x_e x_{e'}) \\ &\quad - \frac{2}{n} (\text{var}(x_e)) - \frac{2}{n} (E(x_e))^2 + \frac{1}{n} \text{var}(x_e) + \\ &\quad \frac{n-1}{n} \text{cov}(x_e, x_{e'}) + [E(x_e)]^2 \end{aligned}$$

$$E(x_e x_{e'}) = \text{cov}(x_e, x_{e'}) + [E(x_e)]^2$$

$$\begin{aligned} E \left\{ (x_e - \bar{x})^2 \right\} &= \text{var}(x_e) - \frac{1}{n} \text{var}(x_e) - \frac{2(n-1)}{n} \left\{ \text{cov}(x_e, x_{e'}) + [E(x_e)]^2 \right\} \\ &\quad - \frac{2}{n} [E(x_e)]^2 + 2 [E(x_e)]^2 + \frac{n-1}{n} \text{cov}(x_e, x_{e'}) \\ &= \frac{n-1}{n} \text{var}(x_e) - \frac{n-1}{n} \text{cov}(x_e, x_{e'}) + \left\{ -\frac{2(n-1)}{n} - \frac{2}{n} + 2 \right\} \\ &\quad [E(x_e)]^2 \end{aligned}$$

$$E \left\{ (x_e - \bar{x})^2 \right\} = \frac{n-1}{n} \left\{ \text{var}(x_e) - \text{cov}(x_e, x_{e'}) \right\}$$

$$\boxed{\frac{\sum_e E \left\{ (x_e - \bar{x})^2 \right\}}{n-1} = \text{var}(x_e) - \text{cov}(x_e, x_{e'})}$$

On a ainsi montré que l'espérance mathématique d'un carré moyen intra-parcellaire est le moitié de l'espérance mathématique du carré de la différence entre deux mesures de deux individus quelconques d'une même parcelle.

Annexe 7.

Espérance mathématique du carré moyen intra panel  
lure

la formule de l'annexe 6 sera utilisée

$$E \left[ \frac{\sum (x_{ijkl} - \bar{x})^2}{\sum n_{ijk} - 1} \right] = \frac{1}{2} E \left[ (x_{ijkl} - x_{ijkl'})^2 \right]$$

$$x_{ijkl} = \mu + G_{ijl} + b_k + f_{ijk} + e_{ijk}$$

$$E \{ G_{ijl}^2 \} = \sigma_G^2$$

$$E \{ G_{ijl} \times G_{ijl'} \} = \text{Cov}(F.S.)$$

$$x_{ijkl} - x_{ijkl'} = f_{ijk} - f_{ijk'} + G_{ijl} - G_{ijl'}$$

$$E \left[ (x_{ijkl} - x_{ijkl'})^2 \right] = E(f_{ijk}^2) + E(f_{ijk'}^2) + E(G_{ijl}^2) + E(G_{ijl'}^2) - 2E(G_{ijl} G_{ijl'})$$

$$\frac{1}{2} E \left\{ (x_{ijkl} - x_{ijkl'})^2 \right\} = \sigma_f^2 + \sigma_G^2 - \text{Cov.}(F.S.)$$

<p>Espérance mathématique du carré moyen intraparcellelure</p> $= \sigma_f^2 + \sigma_G^2 - \text{Cov.}(F.S.)$
--

## Annexe 2. 8

Variance et covariances des moyennes partielles :  
les résultats sont résumés dans le Tableau suivant :

Covariance ou variance	Expression	Nombre de termes.
$y_{ijk}$	$\sigma_2^2 + \sigma_b^2 + \frac{1}{n_{ijk}} \sigma_1^2$	$n p (p-1)$
$y_{ijk}, y_{ijk'}$	$\text{Cov. (FS)}$	$n (n-1) p (p-1)$
$y_{ijk}, y_{ij'k}$	$\text{Cov (HS)} + \sigma_b^2$	$2n p (p-1) (p-2)$
$y_{ijk}, y_{ij'k'}$	$\text{Cov (HS)}$	$2n (n-1) p (p-1) (p-2)$
$y_{ijk}, y_{iik}$	$\text{Cov (HS)} + \sigma_b^2$	$2np (p-1) (p-2)$
$y_{ijk}, y_{iik'}$	$\text{Cov (H.S)}$	$2n (n-1) p (p-1) (p-2)$
$y_{ijk}, y_{jik}$	$\text{Cov (FS)} + \sigma_b^2 - \sigma_n^2$	$n p (p-1)$
$y_{ijk}, y_{jik'}$	$\text{Cov (FS)} - \sigma_n^2$	$n (n-1) p (p-1)$
$y_{ijk}, y_{i'j'k}$	$\sigma_b^2$	$n p (p-1) (p-2) (p-3)$
$y_{ijk}, y_{i'j'k'}$	0	$n (n-1) p (p-1) (p-2) (p-3)$

Annexe 9

Esperances des carrés moyens en fonction du modèle génétique. (modèle II group 3)

1. Calcul de  $E\{CM_b\}$ 

$$\text{var}(\bar{y}_{...}) = \text{var}\left(\sum_{ijk} \frac{1}{n_{ijk}} y_{ijk}\right)$$

$$= \sum_{ijk} \text{var}(y_{ijk}) + \sum \text{cov}(y_{ijk}, y_{i'j'k'})$$

$$= \sum_{ijk} (\sigma_e^2 + \sigma_b^2 + \frac{1}{n_{ijk}} \sigma_1^2) + 2(r-1)p(p-1)$$

$$\text{cov}(FS) + 2rp(p-1)(p-2)(\text{cov}(HS) + \sigma_b^2) + 2r(r-1)p(p-1)(p-2)$$

$$\text{cov}(HS) + 2rp(p-1)(p-2)(\text{cov}(HS) + \sigma_b^2) + 2r(r-1)p(p-1)(p-2)$$

$$\text{cov}(HS) + rp(p-1)(\text{cov}(FS) + \sigma_b^2 - \sigma_2^2) + 2(r-1)p(p-1)$$

$$(\text{cov} FS - \sigma_2^2) + rp(p-1)(p-2)(p-3)\sigma_b^2$$

$$= rp(p-1)\sigma_2^2 + rp(p-1)n_h\sigma_1^2 + rp(p-1)$$

$$(r-1)\text{cov}(FS) + r^2p(p-1)\text{cov}(FS) + 4r^2p(p-1)(p-2)\text{cov}(HS)$$

$$+ rp^2(p-1)^2\sigma_b^2 - r^2p(p-1)\sigma_2^2$$

$$E\{\bar{y}_{...}^2\} = \text{var}(\bar{y}_{...}) + [E\{\bar{y}_{...}\}]^2$$

$$= \text{var}(\bar{y}_{...}) + r^2p^2(p-1)^2\mu^2$$

$$\frac{E\{\bar{y}_{...}^2\}}{rp(p-1)} = \sigma^2 + 2r\text{cov}(FS) + 4r(p-2)\text{cov}(HS)$$

$$+ rp(p-1)\mu^2 + p(p-1)\sigma_b^2 - r\sigma_2^2$$

avec  $n_h = \frac{1}{rp(p-1)} \sum_{ijk} \frac{1}{n_{ijk}}$  et  $\sigma_1^2 = \frac{1}{r}\text{cov}(FS) + n_h\sigma_1^2$

$$\text{var}(Y_{..k}) = \text{var}\left(\sum_{ij}^{20} y_{ijk}\right)$$

$$= \sum_{ij} \text{var}(y_{ijk}) + \sum \text{cov}(y_{ijk}, y_{i'j'k})$$

$$= \sum_{ij} (\sigma_z^2 + \sigma_b^2 + \frac{1}{n_{ijk}} \sigma_1^2) + 2p(p-1)(p-2)$$

$$(\text{cov}(HS) + \sigma_b^2) + 2p(p-1)(p-2)(\text{cov}(HS) + \sigma_b^2) + p(p-1)(\text{cov}(FS) + \sigma_b^2 - \sigma_1^2) + p(p-1)(p-2)(p-3)\sigma_b^2$$

$$= p(p-1)\sigma_z^2 + \sum_{ij} \frac{1}{n_{ijk}} \sigma_1^2 + p^2(p-1)^2 \sigma_b^2$$

$$+ 4p(p-1)(p-2)\text{cov}(HS) + p(p-1)\text{cov}(FS) - p(p-1)\sigma_z^2$$

$$E\{Y_{..k}^2\} = \text{var}(Y_{..k}) + [E\{Y_{..k}\}]^2$$

$$= \text{var}(Y_{..k}) + p^2(p-1)^2 \mu^2$$

$$\sum_k \frac{E\{Y_{..k}^2\}}{p(p-1)} = np(p-1)\mu^2 + n\sigma_z^2 + n n_k \sigma_1^2 + np(p-1)\sigma_b^2 + 4n(p-2)\text{cov}(HS) + 2\text{cov}(FS) - n\sigma_z^2$$

$$E\{S_b\} = \sum_k \frac{E\{Y_{..k}^2\}}{p(p-1)} - \frac{E\{Y_{...}^2\}}{np(p-1)}$$

$$= (n-1)\sigma_z^2 + (n-1)p(p-1)\sigma_b^2$$

$$\boxed{E\{CH_b\} = \sigma_z^2 + p(p-1)\sigma_b^2}$$

2° Calcul de  $E\{CH_g\}$

$$\text{var}(Y_{i..} + Y_{.i.}) = \text{var}(Y_{i..}) + \text{var}(Y_{.i.}) + 2\text{cov}(Y_{i..}, Y_{.i.})$$

$$\text{cov}(Y_{i..}, Y_{.i.}) = \text{cov}\left(\sum_{jk} y_{ijk}, \sum_{j'k'} y_{j'i'k'}\right)$$



$$\text{cov}(Y_{i..}, Y_{.i.}) = n^2 (p-1)(p-2) \text{cov}(HS) + n(p-1)^2 \sigma_b^2 + n^2 (p-1) \text{cov}(FS) - n^2 (p-1) \sigma_z^2$$

$$\text{var}(Y_{i..}) = \text{var}\left(\sum_{jk} Y_{ijk}\right)$$

$$= \sum_{jk} \text{var}(Y_{ijk}) + \sum \text{cov}(Y_{ijk}, Y_{ij'k'})$$

$$= n(p-1) \sigma_2^2 + \left(\sum_{jk} \frac{1}{n_{ijk}}\right) \sigma_1^2 + n(p-1)^2 \sigma_b^2$$

$$+ n^2 (p-1)(p-2) \text{cov}(HS) + n(n-1)(p-1) \text{cov}(FS)$$

$$\text{var}(Y_{i..}) + \text{var}(Y_{.i.}) = 2n(p-1) \sigma_2^2 + \sum_{jk} \left(\frac{1}{n_{ijk}} + \frac{1}{n_{jik}}\right) \sigma_1^2$$

$$+ 2n(p-1)^2 \sigma_b^2 + 2n(n-1)(p-1) \text{cov}(FS) + 2n^2 (p-1)(p-2) \text{cov}(HS)$$

$$E\left\{\frac{(Y_{i..} + Y_{.i.})^2}{2n \mu^2}\right\} = \text{var}(Y_{i..}) + \text{var}(Y_{.i.}) + 2 \text{cov}(Y_{i..}, Y_{.i.}) + 4(p-1)^2$$

$$\sum_i E\left\{\frac{(Y_{i..} + Y_{.i.})^2}{2n(p-2)}\right\} = \frac{2np(p-1)^2}{p-2} \mu^2 + \frac{p(p-1)}{p-2} \sigma_2^2 + \frac{p(p-1)n_R \sigma_1^2}{p-2}$$

$$+ \frac{2p(p-1)^2}{p-2} \sigma_b^2 + \frac{2np(p-1)}{p-2} \text{cov}(FS) + 2np(p-1) \text{cov}(HS)$$

$$- \frac{np(p-1)}{p-2} \sigma_z^2 - \frac{p(p-1)}{p-2} \text{cov}(FS)$$

$$E(S_g) = \sum_i \frac{E\{(Y_{i..} + Y_{.i.})^2\}}{2n(p-2)} - \frac{2E\{Y_{...}^2\}}{np(p-2)}$$

$$= (p-1) \sigma_2^2 + n_R (p-1) \sigma_1^2 - (p-1) \text{cov}(FS)$$

$$+ 2n(p-1) \text{cov}(FS)$$

$$+ 2r(p-1)(p-4) \text{cov}(\bar{H}) - r(p-1) \sigma_r^2$$

$$E\{CH_g\} = \sigma_2^2 + n_h \sigma_1^2 - \text{cov}(F) + 2r \text{cov}(F) \\ + 2r(p-4) \text{cov}(H) - r \sigma_r^2$$

$$E\{CH_g\} = \sigma^2 + 2r \text{cov}(F) + 2r(p-4) \text{cov}(H) - r \sigma_r^2$$

3: bereken de  $E\{CH_b\}$

$$\text{var}(y_{ij.} + y_{ji.}) = \text{var}(y_{ij.}) + \text{var}(y_{ji.}) + 2\text{cov}(y_{ij.}, y_{ji.})$$

$$\text{cov}(y_{ij.}, y_{ji.}) = \text{cov}\left(\sum_k y_{ijk}, \sum_{k'} y_{jik'}\right)$$

$$= \sum_{kk'} \text{cov}(y_{ijk}, y_{jik'})$$

$$= r(r-1)(\text{cov}(F) - \sigma_r^2) + r \text{cov}(F)$$

$$+ r \sigma_b^2 - r \sigma_r^2$$

$$\text{var}(y_{ij.}) = \text{var}\left(\sum_k y_{ijk}\right)$$

$$= \sum_k \text{var}(y_{ijk}) + \sum \text{cov}(y_{ijk}, y_{ijk'})$$

$$= \sum_k \left( \sigma_2^2 + \frac{1}{n_{ijk}} \sigma_1^2 + \sigma_b^2 \right) + r(r-1) \text{cov}(F)$$

$$\text{var}(y_{ij.} + y_{ji.}) = 2r \sigma_2^2 + \sum_k \left( \frac{1}{n_{ijk}} + \frac{1}{n_{jik}} \right) \sigma_1^2$$

$$+ 4r \sigma_b^2 + 4r(r-1) \text{cov}(F) + 2r \text{cov}(F)$$

$$- 2r(r-1) \sigma_r^2 - 2r \sigma_r^2$$

$$E\{(Y_{ij.} + Y_{ji.})^2\} = \text{var}(Y_{ij.} + Y_{ji.}) + 11 n^2 \mu^2$$

$$\sum_{i < j} \frac{E\{(Y_{ij.} + Y_{ji.})^2\}}{2n} = \frac{p(p-1)}{2} \sigma_2^2 + \frac{p(p-1)}{2} n_k \sigma_2^2 + p(p-1)$$

$$\sigma_b^2 + np(p-1)\mu^2 + (n-1)p(p-1)\text{cov}(FS) - \frac{(n-1)p(p-1)}{2} \sigma_n^2 + \frac{p(p-1)}{2}$$

$$\text{cov}(FS) - \frac{p(p-1)}{2} \sigma_n^2$$

$$E\{S_0\} = \sum_{i < j} \frac{E\{(Y_{ij.} + Y_{ji.})^2\}}{2n} - \sum_i \frac{E\{(Y_{i..} + Y_{.i.})^2\}}{2n(p-2)} + \frac{E\{Y_{...}^2\}}{n(p-1)(p-2)}$$

$$= \frac{p(p-3)}{2} \sigma^2 + \frac{p(p-3)}{2} 2n \text{cov}(FS) - \frac{11 np(p-3)}{2}$$

$$\text{cov}(HS) - \frac{np(p-3)}{2} \sigma_n^2$$

$$E\{H_0\} = \sigma^2 + 2n \text{cov}(FS) - 4n \text{cov}(HS) - n \sigma_n^2$$

4.2 calcul de  $E\{H_n\}$

$$\text{var}(Y_{ij.} - Y_{ji.}) = \text{var}(Y_{ij.}) + \text{var}(Y_{ji.}) - 2\text{cov}(Y_{ij.}, Y_{ji.})$$

$$= 2\sigma_2^2 + \sum_k \left( \frac{1}{n_{ijk}} + \frac{1}{n_{jik}} \right) \sigma_1^2 + 2\sigma_b^2$$

$$+ 2n(n-1)\text{cov}(FS) - 2n(n-1)\text{cov}(FS) + 4n(n-1)\sigma_n^2 - 2n\text{cov}(FS)$$

$$- 2n\sigma_b^2 + 4n\sigma_n^2$$

$$E\{(Y_{ij.} - Y_{ji.})^2\} = \text{var}(Y_{ij.} - Y_{ji.})$$

$$E\{S_{\pi}\} = \sum_{i < j} \frac{E\{(y_{ij\cdot} - y_{j\cdot i})^2\}}{2n}$$

$$= \frac{p(p-1)}{2} \sigma^2 + 2n \frac{p(p-1)}{2} \sigma_{\pi}^2$$

$$E\{CH_{\pi}\} = \sigma^2 + 2n \sigma_{\pi}^2$$

H. more 10.

$$\begin{aligned}
 \sigma^2 = & \frac{3}{2} \sigma_A^2 + \frac{29}{48} \sigma_D^2 + \frac{5}{6} \sigma_T^2 + \frac{1}{6} \sigma_Q^2 + \frac{9}{3} \sigma_{HH}^2 + \frac{13}{9} \sigma_{HD}^2 + \frac{1}{6} \sigma_{HT}^2 \\
 & + \frac{16}{6} \sigma_{H\varphi}^2 + \frac{131}{216} \sigma_{DD}^2 + \frac{9}{6} \sigma_{DT}^2 + \frac{1}{6} \sigma_{D\varphi}^2 + \frac{2}{6} \sigma_{TT}^2 + \frac{1}{6} \sigma_{T\varphi}^2 + \frac{1}{6} \sigma_{\varphi\varphi}^2 \\
 & + 2 E \{ \beta_u^2 \} + \frac{16}{3} E \{ \delta_{uy}^2 \} + \frac{1}{3} E \{ \delta_{uyx}^2 \} + \frac{4}{3} E \{ \delta_{uyk}^2 \} \\
 & + 8 E \{ \alpha_i \beta_u \} + \frac{32}{3} E \{ \alpha_j \delta_{uy} \} + \frac{6}{3} E \{ \alpha_k \delta_{uyx} \} + \frac{64}{3} E \{ \beta_y \delta_{uy} \} \\
 & + \frac{2}{3} E \{ \beta_y \delta_{uyx} \} + \frac{1}{3} E \{ \beta_{jk} \delta_{uyk} \} + \frac{8}{3} E \{ \beta_u \delta_{yjk} \} + \frac{4}{3} E \{ \beta_u \delta_{uyj} \} \\
 & + \frac{8}{3} E \{ \delta_{uy} \delta_{yjk} \} + \frac{8}{3} E \{ \delta_{uy} \delta_{uyx} \} + \frac{16}{3} E \{ \delta_{yjk} \delta_{uyk} \} + \frac{1}{3} [E \{ \beta_u \}] \\
 & + \frac{40}{3} E \{ (\alpha\beta)_{i, \ell}^2 \} + \frac{16}{3} E \{ (\alpha\beta)_{i, \ell} (\alpha\beta)_{i, mm} \} + \frac{45}{3} E \{ (\alpha\beta)_{i, \ell} (\alpha\beta)_{i, i} \} \\
 & + \frac{74}{3} E \{ (\alpha\gamma)_{i, i, mm}^2 \} + \frac{32}{3} E \{ (\alpha\gamma)_{i, i, mm} (\alpha\gamma)_{i, m, \ell} \} + \frac{32}{3} E \{ (\alpha\gamma)_{i, i, mm} (\alpha\gamma)_{i, j, mm} \} \\
 & + \frac{5}{3} E \{ (\alpha\delta)_{i, \ell, mm}^2 \} + \frac{16}{3} E \{ (\alpha\delta)_{i, \ell, mm} (\alpha\delta)_{i, \ell, yj} \} + 8 E \{ (\alpha\delta)_{i, \ell, mm}^2 \} \\
 & + \frac{4}{3} E \{ (\beta\beta)_{u, \ell\ell}^2 \} + \frac{62}{3} E \{ (\beta\beta)_{u, i, m}^2 \} + \frac{4}{3} E \{ (\beta\beta)_{u, \ell\ell} (\beta\beta)_{y, \ell\ell} \} \\
 & + \frac{5}{3} E \{ (\beta\beta)_{u, \ell\ell} (\beta\beta)_{y, mm} \} + \frac{23}{3} E \{ (\beta\delta)_{y, i, mm}^2 \} + \frac{196}{3} E \{ (\beta\delta)_{y, i, mm}^2 \} \\
 & + \frac{16}{3} E \{ (\beta\delta)_{y, i, mm}^2 \} + \frac{16}{3} E \{ (\beta\delta)_{y, i, mm} (\beta\delta)_{u, m, \ell\ell} \} \\
 & + \frac{128}{3} E \{ (\beta\delta)_{y, i, mm} (\beta\delta)_{y, m, \ell\ell} \} + \frac{16}{3} E \{ (\beta\delta)_{y, i, mm} (\beta\delta)_{y, j, \ell\ell} \} + \frac{16}{3} E \{ (\beta\delta)_{y, i, mm}^2 \} \\
 & + \frac{2}{3} E \{ (\beta\delta)_{u, \ell, mm}^2 \} + \frac{16}{3} E \{ (\beta\delta)_{y, i, \ell, mm}^2 \} \\
 & + \frac{4}{3} E \{ (\beta\delta)_{u, \ell, mm}^2 \} + 12 E \{ (\beta\delta)_{y, i, \ell, mm}^2 \} + \frac{2}{3} E \{ (\beta\delta)_{u, \ell, mm}^2 \} \\
 & + \frac{4}{3} E \{ (\beta\delta)_{y, i, \ell, mm}^2 \} + \frac{4}{3} E \{ (\beta\delta)_{u, i, mm} (\beta\delta)_{u, m, yj} \} + \frac{16}{3} E \{ (\beta\delta)_{y, i, mm} (\beta\delta)_{y, j, \ell, mm} \} \\
 & + \frac{16}{3} E \{ (\beta\delta)_{y, i, mm} (\beta\delta)_{u, m, yj} \} + \frac{40}{3} E \{ (\gamma\delta)_{y, i, mm}^2 \}
 \end{aligned}$$

$$\begin{aligned}
& + \frac{64}{3} E \{ (\delta\delta)_{\alpha\delta\delta, \ell m n} (\delta\delta)_{\alpha\delta\delta, \ell m} \} + \frac{32}{3} E \{ (\delta\delta)_{\alpha\delta\delta, \ell m} (\delta\delta)_{\alpha\delta\delta, \ell m m} \} \\
& + \frac{32}{3} E \{ (\delta\delta)_{\alpha\delta\delta, \ell m}^2 \} + \frac{8}{3} E \{ (\delta\delta)_{\alpha\delta\delta, \ell m m}^2 \} + \frac{8}{3} E \{ (\delta\delta)_{\alpha\delta\delta, \ell m m}^2 \} \\
& + \frac{16}{3} E \{ (\delta\delta)_{\alpha\delta\delta, \ell m m} (\delta\delta)_{\alpha\delta\delta, \ell m m} \} + \frac{16}{3} E \{ (\delta\delta)_{\alpha\delta\delta, \ell m m}^2 \} \\
& + \frac{8}{3} E \{ (\delta\delta)_{\alpha\delta\delta, \ell m m}^2 \} + \frac{16}{3} E \{ (\delta\delta)_{\alpha\delta\delta, \ell m n} (\delta\delta)_{\alpha\delta\delta, \ell m n, \alpha\delta\delta} \} + \\
& - \frac{1}{6} E \{ (\delta\delta)_{\alpha\delta\delta, \ell m m}^2 \} + \frac{2}{3} E \{ (\delta\delta)_{\alpha\delta\delta, \ell m n}^2 \} - 4 [E \{ \beta_u \}] - \\
& \frac{1}{9} [E \{ \delta_{\alpha\delta\delta} \}]^2 - \frac{16}{9} [E \{ (\beta\beta)_{\alpha\delta\delta, \ell m} \}]^2 - \frac{4}{9} [E \{ (\beta\delta)_{\alpha\delta\delta, \ell m m} \}]^2 \\
& - \frac{1}{36} [E \{ (\delta\delta)_{\alpha\delta\delta, \ell m m} \}]^2 - \frac{4}{3} E \{ \beta_u \} E \{ \delta_{\alpha\delta\delta} \} - \frac{16}{3} E \{ \beta_u \} E \{ (\beta\beta)_{\alpha\delta\delta} \} \\
& - \frac{8}{3} E \{ \beta_u \} E \{ (\beta\delta)_{\alpha\delta\delta, \ell m m} \} - \frac{2}{3} E \{ \beta_u \} E \{ (\delta\delta)_{\alpha\delta\delta, \ell m m} \} \\
& - \frac{8}{9} E \{ \delta_{\alpha\delta\delta} \} E \{ (\beta\beta)_{\alpha\delta\delta, \ell m} \} - \frac{4}{9} E \{ \delta_{\alpha\delta\delta} \} E \{ (\beta\delta)_{\alpha\delta\delta, \ell m m} \} \\
& - \frac{1}{9} E \{ \delta_{\alpha\delta\delta} \} E \{ (\delta\delta)_{\alpha\delta\delta, \ell m m} \} - \frac{16}{9} E \{ (\beta\beta)_{\alpha\delta\delta, \ell m} \} E \{ (\beta\delta)_{\alpha\delta\delta, \ell m m} \} \\
& - \frac{4}{9} E \{ (\beta\beta)_{\alpha\delta\delta, \ell m} \} E \{ (\delta\delta)_{\alpha\delta\delta, \ell m m} \} - \frac{2}{9} E \{ (\beta\delta)_{\alpha\delta\delta, \ell m m} \} E \{ (\delta\delta)_{\alpha\delta\delta, \ell m m} \}
\end{aligned}$$

## Annexe 11.

Formules donnant l'intervalle de confiance des estimations des variances - (Scheffé - 1959)

$\phi$  est estimé par la différence de deux carrés moyens  $\hat{\phi} = MS_1 - MS_2$

$\nu_1, \nu_2$  nombres de degrés de liberté de  $MS_1, MS_2$

$$\text{soit : } \bar{F} = \frac{MS_1}{MS_2}$$

La limite de confiance supérieure de  $\phi$  au risque de  $\alpha_1$  est  $MS_2 g_u(\bar{F})$  avec

$$g_u(\bar{F}) = F_{\alpha_1, \infty, \nu_1} \bar{F} - 1 + \frac{1}{F_{\alpha_1, \nu_2, \nu_1} \bar{F}} \left( 1 - \frac{F_{\alpha_1, \infty, \nu_1}}{F_{\alpha_1, \nu_2, \nu_1}} \right)$$

$$\text{pour } \bar{F} \geq \frac{1}{F_{\alpha_1, \nu_2, \nu_1}}$$

$$g_u(\bar{F}) = 0 \text{ pour } \bar{F} \leq \frac{1}{F_{\alpha_1, \nu_2, \nu_1}}$$

La limite de confiance inférieure de  $\phi$  au risque  $\alpha_2$  est  $MS_2 g_l(\bar{F})$  avec

$$g_l(\bar{F}) = \frac{\bar{F}}{F_{\alpha_2, \nu_1, \infty}} - 1 - \frac{F_{\alpha_2, \nu_1, \nu_2}}{\bar{F}} \left( \frac{F_{\alpha_2, \nu_1, \nu_2}}{F_{\alpha_2, \nu_1, \infty}} - 1 \right)$$

$$\text{pour } \bar{F} \geq F_{\alpha_2, \nu_1, \nu_2}$$

$$\text{et } g_l(\bar{F}) = 0 \text{ pour } \bar{F} \leq F_{\alpha_2, \nu_1, \nu_2}$$

Annexe 1 Moyennes des hybrides et des autofécondations. Caractère : poids en vert en dg  
 Moyenne générale des hybrides = 2,422. Moyenne générale autofécondation = 1,824

♀ ♂	1	2	3	4	5	6	7	8	9	10
1	1,392	3,554	2,525	4,137	3,017	3,750	4,117	2,636	1,950	2,114
2	2,351	3,204	2,716	3,242	2,533	2,371	3,453	3,296	2,345	1,354
3	1,736	2,417	1,239	2,142	2,366	2,151	2,528	2,532	2,045	2,779
4	1,334	3,000	2,732	1,560	2,550	1,542	1,300	2,233	1,315	1,450
5	1,559	2,777	3,261	2,416	2,346	2,374	1,592	2,312	2,237	2,596
6	1,311	2,758	1,570	1,345	2,163	4,674	3,719	1,558	1,573	3,145
7	2,523	3,065	2,237	2,292	2,263	1,541	1,653	2,150	1,279	1,415
8	3,353	3,300	3,229	2,233	3,092	3,442	2,321	1,542	2,139	3,454
9	1,427	2,132	1,746	1,454	1,725	1,775	1,617	2,074	1,108	1,310
10	2,574	3,101	2,429	1,936	2,511	1,577	1,435	2,600	2,330	1,525



Moyennes des hybrides et des autofécondations.

Caractère : Hauteur en cm

Moyenne générale hybrides: 13,24

Moyenne générale autofécondations: 11,26

♀ ♂	1	2	3	4	5	6	7	8	9	10
1	5,84	15,95	11,42	14,76	13,01	14,72	15,98	11,75	8,13	9,32
2	11,67	17,85	16,16	16,51	16,77	15,03	13,80	17,44	15,15	8,92
3	9,06	14,12	8,67	11,73	14,50	16,22	16,47	11,75	11,38	15,54
4	8,15	15,65	13,72	9,45	13,77	11,41	8,71	13,08	7,27	8,29
5	8,60	15,87	17,77	12,77	15,93	15,59	11,64	14,19	12,37	16,36
6	10,49	16,84	11,13	11,34	13,73	13,55	17,69	12,73	9,95	17,19
7	13,14	17,00	13,18	13,80	14,19	12,00	11,34	13,10	8,81	9,23
8	11,82	18,16	17,33	13,38	17,60	18,89	14,60	11,75	11,38	18,53
9	6,96	11,88	10,66	8,21	10,00	11,27	9,99	11,88	7,52	9,69
10	11,94	16,93	14,24	11,93	15,29	11,92	12,51	14,41	12,92	10,29

Moyennes des hybrides et des autofécondations

Moyenne générale hybrides : 6,19

Caractère : Nombre de feuilles.

Moyenne générale autofécondations : 5,55

♀ ♂	1	2	3	4	5	6	7	8	9	10
1	5,06	6,96	6,40	6,63	6,38	6,65	6,43	6,70	5,63	5,33
2	6,23	7,10	6,45	6,48	6,73	6,33	6,61	7,05	6,70	4,41
3	5,38	6,75	6,10	6,11	6,73	6,73	6,96	7,40	6,26	6,33
4	5,33	6,48	6,76	5,36	6,16	5,80	4,86	6,33	5,05	4,95
5	5,10	6,33	7,45	6,22	6,46	6,50	5,61	6,78	6,46	6,25
6	5,98	6,66	6,10	5,98	6,33	6,31	7,45	6,13	5,61	6,33
7	5,31	6,38	6,13	5,86	6,11	5,63	5,20	5,95	4,90	4,66
8	6,36	7,48	7,98	6,33	7,35	7,85	6,51	6,96	6,46	6,61
9	5,06	5,73	5,80	5,28	5,75	5,60	5,36	6,33	5,31	5,10
10	5,70	6,21	6,01	5,28	6,03	5,41	5,21	6,01	5,66	4,93

Analyse de la variance sur les moyennes parcelles.

Caractère : hauteur (en cm)

m : 13,25

Variation	ddl	Somme des carrés	Carré moyen	F
Blees	5	5554		
Grassements	39	4773	53,62921	6,62 (59.445)
Résiduelle	445	3605	8,10112	
Totale	539	13932		

5

2,846.

c.v. 21,48 %.

Variation	ddl	Somme des carrés	Carré moyen	F
Blees	5	5554		
A.G.C	9	2561	284,555	12,48048 (9.35)
A.G.S	35	798	22,800	2,81443 (35.445)
Ef. Rouproques	45	1414	31,422	3,57372 (45.445)
Residu	445	3605	8,10112	
Totale	539	13.932		

Analyse de la variance sur les moyennes parcelles.  
Caractère : Nombre de feuilles.  $n = 6,19$ .

Variation	ddl	Somme des carrés	Carré moyen	F
Blanc	5	175		
Grosses dents	89	274	3,079	9,33 (89.445)
Renduelle	445	147	0,330	
Totale	539	596		

$\sigma = 0,574$

CV: 9,27 %.

Variation	ddl	Somme des carrés	Carré moyen	F
Blanc	5	175		
H.G.C	9	164	18,222	17,23957 (9.35)
H.G.S	35	37	1,057	3,20303 (35.445)
Ef. au poquet	45	73	1,622	4,91515 (45-445)
Rendu	445	147	0,330	
Totale	539	596		

Caractère : hauteur

Intervalle de confiance

	estimation	limite inférieure	limite supérieure
$\hat{\sigma}_g^2$	2,72661	1,15824	9,62988
$\hat{\sigma}_1^2$	1,22492	0,56619	2,56645
$\hat{\sigma}_2^2$	1,94342	0,95125	3,49001

Caractère : Nombre de feuilles.

	estimation	limite inférieure	limite supérieure
$\hat{\sigma}_g^2$	0,17880	0,07864	0,62108
$\hat{\sigma}_1^2$	0,06058	0,03014	0,12275
$\hat{\sigma}_2^2$	0,10767	0,05995	0,18747

Caractère : hauteur

Variance :

	estimation	variance
$\hat{\sigma}_g^2$	2,72661	0,1984671
$\hat{\sigma}_s^2$	1,22492	0,2083333
$\hat{\sigma}_v^2$	1,94342	0,3067833

Caractère : nombre de pailles.

	estimation	variance
$\hat{\sigma}_g^2$	0,17380	0,0050133
$\hat{\sigma}_s^2$	0,06058	0,0004773
$\hat{\sigma}_v^2$	0,10767	0,0012553

Hauter

variation	ddl	Summe des carrés	Carrés moyens	F'
Blues	5	5.554		
H6C	9	2.561	284, 555	7,350 (9.8) S
H5C	35	798	22,800	0,831 NS (35.36)
Maternelle	9	426	47,33	1,72. NS (9.36)
Recherche Reo.	36 45	988 1414	27,44 31,422	3,39. HS (36.445) 3,88. HS
Residuelle	445	3605	8,1012	
Totale	539	13,932		

Nombre de feuilles.

Variations	ddl	Somme des carrés	carrés moyens	F
Blocs	5	175		
A.G.C	9	164	18,222	6,330 <u>2,184</u> (9.9)5
H.S.C	35	37	1,017	0,91 NS (35.36)
Maternelle	9	31	3,444	2,95. HS (9.36)
Rempage.	36	42	1,167	3,54 HS (36.445)
Res.	45	73	1,622	4,84 HS
Résiduelle	445	147	0,330	
Totale	539	596		



hauteur

variation	ddl	Somme des carrés	Carrés moyens.
entre parcelles	539	134,269	23,59290
intra parcelles	4650	109,707	
Totale	5139	243.976	

Nb feuilles

variation	ddl	Somme des carrés	Carrés moyens
entre parcelles	539	5.710	1,21605.
intra parcelles	4647	5.651	
Totale	5.186		

Caractère : hauteur. Annexe 38/8.1 Analyse de variance Clones  
Total des deux coupes.

Clones \ Blocs	1	2	3	4	Total.
1	60,79	56,11	64,30	56,95	238,15
2	78,81	78,05	84,65	82,90	324,41
3	79,45	72,10	78,95	68,85	299,35
4	49,77	31,50	38,46	36,75	156,48
5	69,06	70,83	51,22	60,16	251,27
6	61,41	75,56	66,45	56,58	260,00
7	68,61	63,81	65,15	60,00	257,57
8	73,11	53,05	56,22	58,10	240,48
9	37,74	40,05	45,10	43,05	165,94
10	77,70	81,47	81,21	72,10	312,48
Total	656,45	622,53	631,71	595,44	2506,13

Variation	ddl	Somme des carrés	Carrés moyen	F
Blocs	3	190,4627		
Clones	9	7018,1710	779,842	19,730 H.S
Residuelle	24	948,6236	39,52598	
Totale	36	8157,6633		

$$\sigma = 6,28696$$

$$m = 62,653$$

$$C.V. = 10,03\%$$

Annexe 18-2 <sup>39</sup> Analyse de variance (lines).  
 Caractère : Nombre de pailles.  
 Total des deux coupes.

Blocs Lines	1	2	3	4	Total
1	13,4	12,9	13,7	12,1	52,1
2	15,1	13,9	14,3	14,2	57,5
3	15,6	15,5	15,1	14,2	60,4
4	11,7	8,5	10,0	8,8	39,0
5	16,1	13,5	11,6	13,0	54,2
6	13,5	14,3	13,5	11,7	53,0
7	13,5	12,7	13,0	13,2	52,4
8	14,9	13,1	13,3	13,1	54,4
9	10,8	11,6	11,2	10,9	44,5
10	13,3	13,2	12,8	12,2	51,5
Total.	137,9	129,2	128,5	123,4	519,0

Variation	ddl	Somme des carrés	Carrés moyens	F
Blocs	3	10,361		
Lines	9	84,495	9,388	13,301
Residuelle	24	16,939	0,705792	H S
Totale	36	112,295		

↓ 39-3 années manquantes.  
 $\sigma = 0,840$        $m = 12,975$        $CV = 6,47\%$

Clones \ caract.	poros	Hauteur	Nb feuilles
1	1,361	29,769	6,512
2	2,321	40,551	7,197
3	1,929	37,419	7,550
4	0,517	19,560	4,875
5	1,540	31,409	6,775
6	1,414	32,500	6,625
7	1,534	32,196	6,650
8	1,321	30,060	6,800
9	0,674	20,595	5,562
10	1,346	39,060	6,437

Moyennes des descendance  
d'autofécondations.

Caract <sup>n° P2</sup>	1	2	3	4	5	6	7	8	9	10
Poids	1,2137	3,0535	1,1964	1,1881	2,3488	1,5000	1,7416	1,8928	1,3392	1,7917
Hauteur	5,653	17,201	8,131	9,353	16,236	12,941	11,920	11,643	7,429	10,062
Nb feuilles	4,975	7,167	5,993	5,380	6,523	6,095	5,405	6,475	5,213	4,882

Annexe 21-1

42

Hautefrédérations - hauteurs.

Variation	dell	Somme des carrés	Carré moyen	F
Blous	5	212,66		
plantes	9	733,49	81,499	10,26
Pescheville	45	357,41	7,942	
Totale	59	1303,56		

m : 11,112

CV : 25,36 %

σ : 2,8181

Variation	dell	Somme des carrés	Carré moyen	
entre parcelles	59	9123,721		
intra parcelle	360	8261,732	22,94925	
Totale	419	17385,453		

43

Annexe 21-2 Hétéroindulations. Nb de feuilles

Variations	ddl	Somme des carrés	Carré moyen	F
Bloes	5	8,7745		
plantes	9	36,3655	4,0406	9,80
rendu de	45	18,5554	0,4123	
Totale	59	63,6954		

$m = 5,36$

$C.V. = 10,96\%$

$\sigma = 0,6421$

Variation	ddl	Somme des carrés	Carré moyen
entre parcelles	59	445,9905	
intra parcelle	360	450,0000	1,2500
Totale	419	895,9905	